

80049

STIC-Biotech/ChemLib

From: Schmidt, Mary
Sent: Wednesday, November 13, 2002 11:08 AM
To: STIC-Biotech/ChemLib
Subject: sequence search request 09/296,264

Please search SEQ ID NO:33 as follows:

- (1) full-length search (2772 bases),
- (2) fragment search of SEQ ID NO:33, size limit results to 7-100 bases in length.

Thanks,
Melissa
mailboxes 11e12
room 11d05

Edward Hart
Technical Info. Specialist
TIC/Biotech
Tel: 305-9203

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 2002, 00:05:56 : Search time 4809 Seconds
(without alignments)
16775.416 Million cell updates/sec

Title: US-09-296-264-33

Perfect score: 2772

Sequence: 1 atggagaggggctgcgcgcgt.....gtacttattcgaggcatga 2772

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 994094

Minimum DB seq length: 7

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_em.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_or.*
21: em_ov.*
22: em_ph.*
23: em_pi.*
24: em_pi.*
25: em_pi.*
26: em_pi.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	31.6	1.1	61	9	S71567	S71567 catechol O-
2	25.4	0.9	62	9	HS238518	AJ238518 Homo sapi
3	25	0.9	67	6	AX233509	AX233509 Sequence
C 4	24.8	0.9	90	4	S79940S2	S79941 SLC2A1-gluc
C 5	24.8	0.9	90	9	HSEDA12	AF060788 Homo sapi
6	24.8	0.9	100	11	AF235063	AF235063 Mus muscu
7	24.4	0.9	98	11	MMST56	Z36555 M.musculus
C 8	24.2	0.9	72	4	OABAC6S03	U77044 Ovis aries
C 9	24	0.9	24	51	AX156995	U156995 Sequence
10	24	0.9	81	4	OCMY01	V00886 Rabbit myos
11	23.8	0.9	85	9	HSU33792	U33792 Human immat
12	23.6	0.9	76	9	H006874S07	S75133 Homo sapien
C 13	23.4	0.8	97	1	AF223423	AF223423 pseudomon
14	23.4	0.8	99	8	AF324786	AF324786 Pinus ta
C 15	23.2	0.8	67	9	S75425S04	S75432 collagen ty
16	23.2	0.8	84	9	HSU35685	U35685 Human NF1 e
17	23.2	0.8	85	14	MSVTSGSJ	K02857 Murine sarc
18	23.2	0.8	95	9	HOMONEU1	AF011746 Homo sapi
19	23.2	0.8	99	6	AX088739	AX088739 Sequence
C 20	23	0.8	75	1	BSALA	X73846 Bordetella
21	23	0.8	80	6	AX233560	AX233560 Sequence
22	23	0.8	87	6	E27048	E27048 Novel recep
23	23	0.8	99	6	AX015191	AX015191 Sequence
24	23	0.8	100	6	AX210977	AX210977 Sequence
C 25	22.8	0.8	51	6	AX115029	AX115029 Sequence
C 26	22.8	0.8	67	10	MMU403162	AJ403162 M.musculu
C 27	22.8	0.8	75	1	MCTRC	X16746 Mycoplasma
C 28	22.8	0.8	87	11	HUMSWX1900	L42717 Human chrom
29	22.6	0.8	66	4	BTRP5	V00129 Bovine sate
30	22.6	0.8	81	6	AX014760	AX014760 Sequence
C 31	22.6	0.8	85	3	AF318498	AF318498 Scutigere
32	22.6	0.8	88	8	PHVMTLA	M35722 Phageolus v
C 33	22.6	0.8	93	10	AF041922	AF041922 Mus muscu
34	22.6	0.8	95	10	MMV5LK1	Z12421 M.musculus
35	22.6	0.8	97	6	AR140888	AR140888 Sequence
36	22.6	0.8	97	6	AR150838	AR150838 Sequence
37	22.6	0.8	97	6	I65716	I65716 Sequence 76
38	22.6	0.8	97	6	I67948	I67948 Sequence 76
39	22.6	0.8	97	6	I90169	I90169 Sequence 76
40	22.6	0.8	99	4	BTU32912	U32912 Bos taurus
41	22.4	0.8	40	6	AR031807	AR031807 Sequence
42	22.4	0.8	40	6	I24873	I24873 Sequence 1
C 43	22.4	0.8	47	6	AX194712	AX194712 Sequence
44	22.4	0.8	51	6	AX158190	AX158190 Sequence
45	22.4	0.8	60	6	AR031477	AR031477 Sequence

ALIGNMENTS

RESULT 1
S71567
LOCUS
DEFINITION
S71567
c catechol O-methyltransferase [3', region, ribosomal frameshift by
extended C-terminus] [human, Genomic Recombinant, 61 nt].
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

61 bp DNA linear PRI 13-OCT-2000
S71567
c catechol O-methyltransferase [3', region, ribosomal frameshift by
extended C-terminus] [human, Genomic Recombinant, 61 nt].
S71567
S71567.1 GI:565175
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 61)
Vilbois,F., Caspers,P., da Prada,M., Lang,G., Karrer,C., Lahm,H.W.
and Cesura,A.M.

Pred. No. is the number of results predicted by chance to have a

[illegible]


```

Db      62 CTCA 65

RESULT 11
HSU33792
LOCUS      HSU33792      85 bp      mRNA      linear      PRI 25-MAR-1999
DEFINITION Human immature B cell Ig heavy chain mRNA, third
              complementarity-determining region, clone MBT-149, partial cds.
ACCESSION  U33792
VERSION     U33792.1 GI:4096621
KEYWORDS
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 85)
AUTHORS     Raaporst,F.M., Tami,J., Fischbach,M. and Sanz,I.
TITLE       Direct Submission
JOURNAL     Submitted (14-AUG-1995) Frank Martin Raaporst,
              Medicine/Rheumatology, University of Texas Health and Science
              Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX
              78284-7868, USA
FEATURES    Location/Qualifiers
             source          1..85
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="14"
                        /map="14q32.3"
                        /clone="MBT-149"
                        /cell_type="bone marrow immature B lymphocyte"
                        /dev_stage="adult"
                        <1..>85
             CDS             /note="third complementarity-determining region"
                        /codon_start=1
                        /product="Ig heavy chain"
                        /protein_id="AAC99972.1"
                        /db_xref="GI:4096622"
                        /translation="TAVYICARRYSGSSHTDYWGQGLVTVS"
BASE COUNT  17 a 24 c 25 g 19 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85

Query Match      0.98; Score 23.8; DB 9; Length 85;
Best Local Similarity 59.7%; Pred. No. 5.6e+05;
Matches 40; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 2628 GGTGCTGTGACGTGCTTGGCATAATGGGATGTCAGAAAGAACTTGTCTGCCCT 2687
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 GCGCGTGATTACTGTGGAGACGTTATAGTGGGAGCTCACACACTGACTACTGGGGCCA 62
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 2688 GGAGAAC 2694
      |||||
Db 63 GGGAAACC 69

RESULT 12
H006874S07
LOCUS      H006874S07      76 bp      DNA      linear      PRI 29-JAN-2001
DEFINITION Homo sapiens megakaryocyte-associated tyrosine kinase (MATK) gene,
              exon 7.
ACCESSION  S75153
VERSION     S75153.1 GI:896211
KEYWORDS
SEGMENT
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 76)
AUTHORS     Avraham,S., Jiang,S., Ota,S., Fu,Y., Deng,B., Dowler,L.L.,
              White,R.A. and Avraham,H.
TITLE       Structural and functional studies of the intracellular tyrosine
              kinase MATK gene and its translated product
JOURNAL     J. Biol. Chem. 270 (4), 1833-1842 (1995)

MEDLINE 95130565
PUBMED   7530249
REMARK   GenBank staff at the National Library of Medicine created this
          entry [NCBI gibbsq 159429] from the original journal article.
          This sequence comes from Fig. 5.
FEATURES    Location/Qualifiers
             source          1..76
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
             exon           7..72
                        /gene="MATK"
                        /number=7
BASE COUNT  20 a 14 c 27 g 15 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76

Query Match      0.98; Score 23.6; DB 9; Length 76;
Best Local Similarity 61.3%; Pred. No. 6.2e+05;
Matches 38; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 2316 GTCCTCGAAACTTTATCAGGTGATTTTCGAGGGCGAAATCGAAAGAAACCTTGGTGG 2375
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 14 GTTACTGAACCTCGAGCATTTGACATTGGGAGCACAGATCGGAGAGAGATTGGAGG 73
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 2376 GA 2377
      |
Db 74 TA 75

RESULT 13
AF223423/c
LOCUS      AF223423      97 bp      DNA      linear      BCT 29-MAR-2000
DEFINITION Pseudomonas aeruginosa clone 570 biofilm-dependent regulatory
              element.
ACCESSION  AF223423
VERSION     AF223423.1 GI:7340077
KEYWORDS
SOURCE      Pseudomonas aeruginosa.
ORGANISM    Pseudomonas aeruginosa
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Pseudomonas.
REFERENCE   1 (bases 1 to 97)
AUTHORS     Weyers,J.B., van Zyl,W.H. and Brozel,V.S.
TITLE       Pseudomonas aeruginosa contains a novel and diverse group of
              biofilm-dependent regulatory elements
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 97)
AUTHORS     Weyers,J.B. and Brozel,V.S.
TITLE       Direct Submission
JOURNAL     Submitted (11-JAN-2000) Microbiology and Plant Pathology,
              University of Pretoria, Lunnon Road, Pretoria, Gauteng 0002, South
              Africa
FEATURES    Location/Qualifiers
             source          1..97
                        /organism="Pseudomonas aeruginosa"
                        /strain="PAOI"
                        /db_xref="taxon:287"
                        /clone="570"
             misc_feature     1..97
                        /note="biofilm-dependent regulatory element; up-regulated
                        when grown as a biofilm on a solid surface"
BASE COUNT  16 a 35 c 35 g 11 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97

Query Match      0.88; Score 23.4; DB 1; Length 97;
Best Local Similarity 57.5%; Pred. No. 7.2e+05;
Matches 42; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 5 AGAGGGGCTCCGCTCTCTGCGCGCGTCTGCGCCCTCGTCTCTCGCCCGCGCGGCTT 64
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75 AGGTGAACAGCGCCGCTCGCGCTGCTGCGACGCGACCTCGATGTCGCGCGGCTTGGCT 16
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 65 TTCGCAACGATGA 77
      |. | |||||

```

Db 15 TCCTCGACGATGA 3

RESULT 14
AF324786
LOCUS AF324786 99 bp DNA linear PLN 10-MAY-2002
DEFINITION Pinus taeda clone PTrX4115 microsatellite sequence.
ACCESSION AF324786
VERSION AF324786.1 GI:14029429
KEYWORDS
SOURCE Pinus taeda.
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 99)
REFERENCE Zhou, Y., Bui, T., Auckland, L.D. and Williams, C.G.
AUTHORS Undermethylated DNA as a source of microsatellites from a conifer genome
TITLE Undermethylated DNA as a source of microsatellites from a conifer genome
JOURNAL Genome 45 (1), 91-99 (2002)
MEDLINE 21905334
PUBMED 11908673
REFERENCE 2 (bases 1 to 99)
AUTHORS Zhou, Y., Auckland, L.D. and Williams, C.G.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-2000) Forest Science Department, Texas A&M University, Room 317 HFSB, College Station, TX 77843, USA

FEATURES
source
1..99
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="PTrX4115"
repeat_region
1..99
/note="microsatellite"
/rpt_type=tandem

BASE COUNT 30 a 24 c 10 g 35 t
ORIGIN

Query Match 0.8%; Score 23.4; DB 8; Length 99;
Best Local Similarity 60.0%; Pred. No. 7.3e+05;
Matches 39; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 2538 GCCAGCATGTGTTGACGACCTAGAACCCATCCTCATCACCATGATGCCATGAGCGC 2597
Db 28 GACAATAAATTTTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCGA 87

QY 2598 CTTGG 2602
Db 88 TCTGG 92

RESULT 15
S75425S04/c
LOCUS S75425S04 67 bp DNA linear PRI 02-JUN-2000
DEFINITION collagen type VI alpha 2(VI) chain [exons 1-19 and intron/exon junctions] [human, Genomic, 67 nt, segment 4 of 19].
ACCESSION S75432
VERSION S75432.1 GI:241989
KEYWORDS
SEGMENT 4 of 19
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 67)
AUTHORS Saitta, B., Wang, Y.M., Renkart, L., Zhang, R.Z., Pan, T.C., Timpl, R. and Chu, M.L.
TITLE The exon organization of the triple-helical coding regions of the human alpha 1(VI) and alpha 2(VI) collagen genes is highly similar
JOURNAL Genomics 11 (1), 145-153 (1991)
MEDLINE 92112205
PUBMED 1765372
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 75432] from the original journal article.

This sequence comes from Figure 4.
Map location: chromosome 21.
FEATURES
Source
1..67
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
31..57
/number=4
BASE COUNT 12 a 23 c 17 g 15 t
ORIGIN

Query Match 0.8%; Score 23.2; DB 9; Length 67;
Best Local Similarity 70.5%; Pred. No. 7.9e+05;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2562 AGAACCATCTCTCATCACCATCATGACCATGAGCGCCCTGGGG 2605
Db 67 AGAGCCTCACCTTCTCTCTTTTGAGCCAGGAACGCCCTGGGG 24

Search completed: November 20, 2002, 02:37:56
Job time : 4822 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2002, 20:48:10 ; Search time 384 Seconds
(without alignments)
16256.613 Million cell updates/sec

Title: US-09-296-264-33
Perfect score: 2772
Sequence: 1 atgagagggggtgcgcgt.....gtacttattcggaggcatga 2772

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 2398870

Minimum DB seq length: 7
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				N_Geneseq_101002.*
1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*			
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*			
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*			
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*			
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*			
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*			
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*			
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*			
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*			
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*			
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*			
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*			
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*			
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*			
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*			
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*			
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*			
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*			
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*			
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*			
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*			
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*			
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*			
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	60	2.2	60 24	ABN43139 Human spliced tran
2	50.6	1.8	65 24	ABN56218 Mouse spliced tran
3	30	1.1	30 22	AAF83316 Human neuropilin-1
C 4	27	1.0	27 22	AAF83315 Human neuropilin-1
C 5	26.6	1.0	96 4	AA30181 Sequence of synthe
C 6	25.4	0.9	51 22	AA227936 Human SNP oligonuc
C 7	25.4	0.9	65 24	ABN55423 Mouse spliced tran
C 8	25	0.9	25 24	ABA98777 Neuropiline marker
9	25	0.9	67 22	AA543690 Corneodesmosin sin

C	10	25	0.9	96	4	AA30180	Sequence of synthe
	11	24.8	0.9	65	24	ABN58419	Mouse spliced tran
	12	24.6	0.9	93	24	ABN70179	Streptococcus poly
	13	24.6	0.9	95	16	AAT23132	Human gene signatu
	14	24.4	0.9	90	22	ABA48667	Human breast cell
	15	24.4	0.9	90	22	ABA66579	Human foetal liver
	16	24.4	0.9	90	22	ABA33646	Probe #12112 for g
	17	24.4	0.9	90	22	AAK15006	Human brain expres
	18	24.4	0.9	90	22	AAK40737	Human bone marrow
	19	24.4	0.9	90	22	AAI21500	Probe #11433 for g
	20	24.4	0.9	90	22	AAI46792	Probe #15478 used
	21	24.4	0.9	90	22	AAI07197	Probe #7188 used t
	22	24.4	0.9	90	24	ABS14693	Human genome-deriv
C	23	24.2	0.9	80	21	ABL50868	Modified erythrope
	24	24.2	0.9	87	21	AAAS9449	Oligonucleotide us
	25	24	0.9	24	24	ABA98776	Neuropiline marker
C	26	24	0.9	34	22	AAF83317	Human neuropilin-1
	27	24	0.9	51	22	AAI73382	Human silent SNP c
C	28	24	0.9	96	20	AA211484	Sequence 2.5-5 ide
	29	23.8	0.9	59	22	AAH26556	Kidney cancer spec
C	30	23.8	0.9	61	21	AAI1326	Human secreted pro
C	31	23.8	0.9	98	22	ABA48465	Human breast cell
C	32	23.8	0.9	98	22	ABA68370	Human foetal liver
C	33	23.8	0.9	98	22	ABA69527	Human foetal liver
C	34	23.8	0.9	98	22	ABA33432	Probe #11898 for g
C	35	23.8	0.9	98	22	ABA36451	Probe #14917 for g
C	36	23.8	0.9	98	22	AAK14787	Human brain expres
C	37	23.8	0.9	98	22	AAK40528	Human bone marrow
C	38	23.8	0.9	98	22	AAI21287	Probe #11220 for g
C	39	23.8	0.9	98	22	AAI46563	Probe #15249 used
C	40	23.8	0.9	98	22	AAI49652	Probe #18338 used
C	41	23.8	0.9	98	22	AAI05991	Probe #6982 used t
C	42	23.8	0.9	98	22	AAI09921	Probe #9912 used t
C	43	23.8	0.9	98	24	ABS14504	Human genome-deriv
C	44	23.8	0.9	98	24	ABS17765	Human genome-deriv
C	45	23.6	0.9	65	24	ABN28748	Rat spliced trans

ALIGNMENTS

RESULT 1
ABN43139
ID ABN43139 standard; DNA; 60 BP.
XX
XX
AC ABN43139;
XX
DT 15-JUL-2002 (first entry)
DE Human spliced transcript detection oligonucleotide SEQ ID NO:15887.
XX
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
OS Homo sapiens.
PN WO200210449-A3
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-241607P.
PR 02-MAY-2001; 2001US-287924P.
(COMP-) COMPUGEN INC.
Shoshan A Wasserman A, Mintz E, Mintz L, Faigler S;
WPI; 2002-257382730.
New oligonucleotide libraries comprising oligonucleotides which
selectively hybridize to mRNAs transcribed from a transcription unit of

PT a genome, useful for detecting tissue-, pathology-, and
 XX developmental-specific genes -
 PS Example 1; SEQ ID 15887; 47pp; English.

CC The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridising selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 60 BP; 18 A; 9 C; 22 G; 11 T; 0 other;

Query Match 2.2%; Score 60; DB 24; Length 60;
 Best Local Similarity 100.0%; Pred. No. 2.4e-08;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1484 TGAGGGCATCATCTTCAGGGTGGGAGACCCGAGAGACAGGTGTTGATGAGGAAGT 1543
 Db 1 TGAGGGCATCATCTTCAGGGTGGGAGACCCGAGAGACAGGTGTTGATGAGGAAGT 60

RESULT 2

ABN56218
 ID ABN56218 standard; DNA; 65 BP.

XX AC ABN56218;

XX DT 15-JUL-2002 (first entry)

XX DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:28966.

XX DE Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.

XX OS Mus musculus.

XX PN WO200210449-A2.

XX PD 07-FEB-2002.

XX PF 20-JUL-2001; 2001WO-1B01903.

XX PR 28-JUL-2000; 2000US-221607P.

XX PR 02-MAY-2001; 2001US-287724P.

XX PA (COMP-) COMPUGEN INC.

XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Paigler S;

XX DR WPI; 2002-257383/30.

XX PT New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of

PT a genome, useful for detecting tissue-, pathology-, and
 XX developmental-specific genes -
 PS Example 1; SEQ ID 28966; 47pp; English.

CC The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridising selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 65 BP; 16 A; 21 C; 12 G; 16 T; 0 other;

Query Match 1.8%; Score 50.6; DB 24; Length 65;
 Best Local Similarity 86.2%; Pred. No. 2.2e-05;
 Matches 56; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1640 ATACACCTGAGCTCGGAGCTTTCCAGCTCTCTCCACGCGATTTCATCAGGATCTACCCCG 1699
 Db 1 ACACACCTGAGCTCGGAGCTTTCCAGCTCTCTCCACGCGATTTCATCAGGATCTACCCCG 60

QY 1700 AGAGA 1704

Db 61 AGAGA 65

RESULT 3

AAF83316

ID AAF83316 standard; DNA; 30 BP.

XX AC AAF83316;

XX DT 09-JUL-2001 (first entry)

XX DE Human neuropilin-1 (NP-1) DNA amplifying sense primer BUL-259.

XX DE Receptor protein; vascular endothelial growth factor receptor-2; NP-1;
 KW VEGFR-2; neuropilin-1; co-receptor; human; angiogenic; PCR primer; ss.

XX OS Homo sapiens.

XX PN WO200131346-A2.

XX PD 03-MAY-2001.

XX PF 26-OCT-2000; 2000WO-US29579.

XX PR 28-OCT-1999; 9906-0162367.

XX PA (PROC) PROCTER & GAMBLE CO.

XX PI Rosenbaum JS, Whitaker GB, Limberg BJ;

XX DR WPI; 2001-308686/32.

XX Determining compounds which bind to a complex comprising vascular
PT endothelial growth factor receptor-2 and Neuropilin-1 to provide
PT superior pro- and anti-angiogenic agents
XX Example 1; Fig 3; 82pp; English.
XX The invention relates to determining whether a compound is capable of
CC binding to a receptor protein complex comprising a vascular endothelial
CC growth factor receptor-2 (VEGFR-2) receptor protein and a neuropilin-1
CC (NP-1) receptor protein. One method comprises introducing a sample
CC comprising the compound to the receptor protein and allowing the
CC compound to bind to the complex. Signaling through VEGFR-2 is enhanced
CC in the presence of the NP-1 co-receptor. The methods of the invention
CC can be used for identifying novel pro- and anti-angiogenic compounds.
CC Sequences AAF83314-16 represents PCR primers used for isolating human
CC NP-1 DNA.
XX Sequence 30 BP; 5 A; 11 C; 8 G; 6 T; 0 other;
SQ Query Match 1.1%; Score 30; DB 22; Length 30;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 901 TGGTCTGCAGAGCGCTCCCGCTGAACCTAC 930
Db 1 TGGTCTGCAGAGCGCTCCCGCTGAACCTAC 30
RESULT 4
AAF83315/c
ID AAF83315 standard; DNA; 27 BP.
XX AC AAF83315;
XX DT 09-JUL-2001 (first entry)
XX Human neuropilin-1 (NP-1) DNA amplifying antisense primer BJL-238R.
XX Receptor protein; vascular endothelial growth factor receptor-2; NP-1;
KW VEGFR-2; neuropilin-1; co-receptor; human; angiogenic; PCR primer; ss.
XX Homo sapiens.
XX WO200131346-A2.
XX 03-MAY-2001.
XX 26-OCT-2000; 2000WO-US29579.
XX 28-OCT-1999; 99US-0162367.
XX (PROC) PROCTER & GAMBLE CO.
XX Rosenbaum JS, Whitaker GB, Limberg RJ;
XX WPI; 2001-308686/32.
XX Determining compounds which bind to a complex comprising vascular
PT endothelial growth factor receptor-2 and Neuropilin-1 to provide
PT superior pro- and anti-angiogenic agents
XX Example 1; Fig 3; 82pp; English.
XX The invention relates to determining whether a compound is capable of
CC binding to a receptor protein complex comprising a vascular endothelial
CC growth factor receptor-2 (VEGFR-2) receptor protein and a neuropilin-1
CC (NP-1) receptor protein. One method comprises introducing a sample
CC comprising the compound to the receptor protein and allowing the
CC compound to bind to the complex. Signaling through VEGFR-2 is enhanced
CC in the presence of the NP-1 co-receptor. The methods of the invention
CC can be used for identifying novel pro- and anti-angiogenic compounds.
CC Sequences AAF83314-16 represents PCR primers used for isolating human

CC NP-1 DNA.
XX Sequence 27 BP; 5 A; 7 C; 8 G; 7 T; 0 other;
SQ Query Match 1.0%; Score 27; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 890 ATAGCACCACCACTGGTCTGCAGAGCGCT 916
Db 27 ATAGCACCACCACTGGTCTGCAGAGCGCT 1
RESULT 5
AAN30181/c
ID AAN30181 standard; DNA; 96 BP.
XX AC AAN30181;
XX DT 05-APR-1992 (first entry)
XX Sequence of synthetic gene encoding human calcitonin analogue.
XX Paget's disease; therapy; hypercalcaemia; synthetic gene;
KW calcitonin analogue; ds.
XX Homo sapiens.
XX WO8304028-A.
XX 24-NOV-1983.
XX 15-APR-1983; 83WO-0901772.
XX 06-MAY-1982; 82US-0375499.
XX (MOLE-) APPL MOLECULAR GENE.
XX (AMGE-) AMGEN INC.
XX Lau E, Suggs S;
XX WPI; 1983-833187/48.
XX Manufactured genes for transformation of Escherichia coli host -
PT for prodn. of human calcitonin poly:peptide or analogue
XX Claim 10; Page 27; 33pp; English.
XX The inventors claim synthetic genes for human calcitonin analogues.
CC The genes may include initial and/or terminal codons specifying
CC additional amino acid(s), and may be fused to a second gene to give a
CC fused polypeptide. A typical analogue prep. is val 8-human
CC calcitonin. The analogues may have high potency and duration of
CC action for use in the treatment of Paget's disease and hypercalcaemia
CC without the prodn. of immune reactions (as can occur when salmon
CC calcitonin is used).
XX Sequence 96 BP; 18 A; 31 C; 22 G; 25 T; 0 other;
SQ Query Match 1.0%; Score 26.6; DB 4; Length 96;
Best Local Similarity 63.1%; Pred. No. 9.6e+02;
Matches 41; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 2528 TCTCAGAACCCAGGCCTGTTGAAGACCTTAGAACCCCTCTCATCACCATCATAG 2587
Db 73 TCTCGCGAAGGTATGGAATTTGTTGAAGTCTGTAGTATAGTGCACGACGAGGTAG 14
Qy 2588 CCATG 2592
Db 13 ACAGG 9
RESULT 6

```

AAL27936
ID AAL27936 standard; DNA; 51 BP.
AC
XX
AC AAL27936;
XX
AC
XX
DT 24-JAN-2002 (first entry)
XX
XX
DE Human SNP oligonucleotide #1144.
XX
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; kinase; colony stimulating factor;
KW complement related protein; cytochrome; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
XX
OS Homo sapiens.
XX
OS
XX
PN WO200147944-A2.
XX
PN
XX
PD 05-JUL-2001.
XX
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
XX
PR 28-DEC-1999; 99US-0173419.
XX
PR 27-DEC-2000; 2000US-0173419.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX
XX
DR WPI; 2001-465210/50.
XX
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
XX
PS Claim 1; Page 1707; 4143pp; English.
XX
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
XX
SQ Sequence 51 BP; 11 A; 18 C; 12 G; 10 T; 0 other;

Query Match 0.9%; Score 25.4; DB 22; Length 51;
Best Local Similarity 74.4%; Pred. No. 1.5e+03;
Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1851 CTTCCAGCTCAGAGTGGCCACACGTGCTGGCCACAGAAAG 1893
DB 9 CTGGAGCTCAGAGTGGCCCCACCTGTACCGCCCTCTGATAAG 51

RESULT 7
ABN55423/C
ID ABN55423 standard; DNA; 65 BP.
XX
XX
AC
XX
DT 24-JAN-2002 (first entry)
XX
XX
DE Human SNP oligonucleotide #1144.
XX
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; kinase; colony stimulating factor;
KW complement related protein; cytochrome; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
XX
OS Homo sapiens.
XX
OS
XX
PN WO200147944-A2.
XX
PN
XX
PD 05-JUL-2001.
XX
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
XX
PR 28-DEC-1999; 99US-0173419.
XX
PR 27-DEC-2000; 2000US-0173419.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX
XX
DR WPI; 2001-465210/50.
XX
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
XX
PS Claim 1; Page 1707; 4143pp; English.
XX
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
XX
SQ Sequence 51 BP; 11 A; 18 C; 12 G; 10 T; 0 other;

Query Match 0.9%; Score 25.4; DB 22; Length 51;
Best Local Similarity 74.4%; Pred. No. 1.5e+03;
Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1851 CTTCCAGCTCAGAGTGGCCACACGTGCTGGCCACAGAAAG 1893
DB 9 CTGGAGCTCAGAGTGGCCCCACCTGTACCGCCCTCTGATAAG 51

RESULT 7
ABN55423/C
ID ABN55423 standard; DNA; 65 BP.
XX
XX
AC
XX
DT 24-JAN-2002 (first entry)
XX
XX
DE Human SNP oligonucleotide #1144.
XX
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; kinase; colony stimulating factor;
KW complement related protein; cytochrome; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
XX
OS Homo sapiens.
XX
OS
XX
PN WO200147944-A2.
XX
PN
XX
PD 05-JUL-2001.
XX
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
XX
PR 28-DEC-1999; 99US-0173419.
XX
PR 27-DEC-2000; 2000US-0173419.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX
XX
DR WPI; 2001-465210/50.
XX
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
XX
PS Claim 1; Page 1707; 4143pp; English.
XX
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiotensin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
XX
SQ Sequence 51 BP; 11 A; 18 C; 12 G; 10 T; 0 other;

Query Match 0.9%; Score 25.4; DB 24; Length 65;
Best Local Similarity 64.4%; Pred. No. 1.8e+03;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 635 AAATCTGGGATGATCCCTGATGTTGGCCCTCAGATTGGCGTGTACTGTGCACAGAA 693
DB 61 AAACACGGACTGGAAACACAGGTGCTGGACATGACACTGGCCGTGACCCGTTGAGAA 3

RESULT 8
ABA98777/C
ID ABA98777 standard; DNA; 25 BP.
XX
XX
AC
XX
DT 15-JUL-2002 (first entry)
XX
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:28171.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
XX
PN WO200210449-A2.
XX
XX
PD 07-FEB-2002.
XX
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
XX
PR 28-JUL-2000; 2000US-221607P.
XX
PR 02-MAY-2001; 2001US-287724P.
XX
XX
PA (COMP-) COMPUGEN INC.
XX
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
XX
DR WPI; 2002-257383/30.
XX
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
XX
PS Example 1; SEQ ID 28171; 47pp; English.
XX
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample. In expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 65 BP; 10 A; 19 C; 16 G; 20 T; 0 other;

Query Match 0.9%; Score 25.4; DB 24; Length 65;
Best Local Similarity 64.4%; Pred. No. 1.8e+03;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 635 AAATCTGGGATGATCCCTGATGTTGGCCCTCAGATTGGCGTGTACTGTGCACAGAA 693
DB 61 AAACACGGACTGGAAACACAGGTGCTGGACATGACACTGGCCGTGACCCGTTGAGAA 3

RESULT 8
ABA98777/C
ID ABA98777 standard; DNA; 25 BP.
XX
XX
AC
XX
DT 15-JUL-2002 (first entry)
XX
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:28171.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
XX
PN WO200210449-A2.
XX
XX
PD 07-FEB-2002.
XX
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
XX
PR 28-JUL-2000; 2000US-221607P.
XX
PR 02-MAY-2001; 2001US-287724P.
XX
XX
PA (COMP-) COMPUGEN INC.
XX
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
XX
DR WPI; 2002-257383/30.
XX
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
XX
PS Example 1; SEQ ID 28171; 47pp; English.
XX
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample. In expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 65 BP; 10 A; 19 C; 16 G; 20 T; 0 other;
```


AC ABA98777;
XX
DT 17-MAY-2002 (first entry)
XX
DE Neuropiline marker NRPI PCR primer #2.
XX
KW PCR; primer; diagnosis; prognosis; prostatic cancer; ss.
XX
OS Synthetic.
XX
PN WO200200929-A1.
XX
PD 03-JAN-2002.
XX
PF 22-MAY-2001; 2001WO-FR01566.
XX
PR 27-JUN-2000; 2000FR-0008266.
XX
PA (UROG-) UROGENE.
XX
PI Cussenot O, Latil A, Chantrel-Groussard K;
XX
DR WPI; 2002-130910/17.
XX
PT Diagnosis and prognosis of prostatic cancer, comprises analysis of
PT specific cell markers, antigens and loss of chromosomal alleles, also
PT used for optimizing treatment.
XX
PS Claim 5; Page 33; 46pp; French.
XX
SQ The present invention relates to a method for in vitro diagnosis and/or
CC prognosis of prostatic cancer. The method comprises: (a) separate or
CC simultaneous extraction of total DNA or RNA from a sample of prostatic
CC cells; (b) analysing the cells for specific markers of prostatic origin
CC from the RNA or by direct measurements of cell-specific antigens in urine
CC present in the sample; and (c) screening for chromosomal alterations,
CC from loss of alleles, in the DNA. The present sequence is a PCR primer
CC for a cellular marker, and was used in the method of the invention.
XX
SQ Sequence 25 BP; 6 A; 3 C; 9 G; 7 T; 0 other;
Query Match 0.9%; Score 25; DB 24; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 874 ACAGCTTCTCCAGTATAGCACCA 898
DB 25 ACAGCTTCTCCAGTATAGCACCA 1
RESULT 9
AAS43690
ID AAS43690 standard; DNA; 67 BP.
XX
AC AAS43690;
XX
DT 18-DEC-2001 (first entry)
XX
DE Corneodesmosin single nucleotide polymorphism #48.
XX
KW Human; single nucleotide polymorphism; SNP; PCR primer; antiinflammatory;
KW antipsoriatic; corneodesmosin; inflammatory disease; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200162788-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-GB00795.
XX
PR 23-FEB-2000; 2000GB-0004312.
XX

PA (OXAG-) OXAGEN LTD.
XX
PI Olaveson M, Lench N, Allen M, Tazi-Ahmini R;
XX
DR WPI; 2001-570627/64.
XX
KW Corneodesmosin protein and polynucleotide encoding it, having one or
PT more polymorphisms useful in treating, diagnosing or determining
PT susceptibility to corneodesmosin-mediated diseases, for e.g.
PT inflammatory diseases
XX
PS Disclosure; Page 39; 60pp; English.
XX
CC The invention relates to corneodesmosin protein (I) and nucleic acid
CC (II) encoding the corneodesmosin gene, where the gene comprises a base
CC substitution, deletion or insertion at one or more positions. (I) and
CC (II) are useful for screening for agents for use in prognosis, diagnosis
CC and treatment of individuals having or being susceptible to
CC corneodesmosin-mediated disease, by monitoring the reaction between the
CC molecules and the agents. The nucleotide and amino acid polymorphisms are
CC useful for diagnosing or determining susceptible to corneodesmosin-
CC mediated disease, which facilitates subsequent treatment of the disease
CC for e.g. inflammatory diseases, in particular psoriasis. Fragments of (I)
CC are useful in diagnostic, prognostic or therapeutic methods and as
CC research tools for e.g. in drug screening. (II) is useful as probes or
CC primers for detecting an allele of the polymorphism or in the regulation
CC of corneodesmosin gene. Antibodies which binds to (I) are useful for
CC screening DNA clone libraries for cells secreting the antigen. (II) is
CC useful as a model to investigate the role of corneodesmosin in normal
CC skin function. AAS43492-AAS43749 represent corneodesmosin coding
CC sequences, single nucleotide polymorphisms (SNPs) and PCR primers of the
CC invention.
XX
SQ Sequence 67 BP; 5 A; 23 C; 7 G; 31 T; 1 other;
Query Match 0.9%; Score 25; DB 22; Length 67;
Best Local Similarity 72.1%; Pred. No. 2.4e+03;
Matches 31; Conservative 1; Mismatches 11; Indels 0; Gaps 0;
QY 336 TGTGCTTTCAGGGCCATTTCTTTTATCATCAATTTGCTCTGAC 378
DB 2 TGTCTCTTTCAGGGTCCTTTCTTTTAGACCTATTCTGCTGCC 44
RESULT 10
AAN30180/C
ID AAN30180 standard; DNA; 96 BP.
XX
AC AAN30180;
XX
DT 05-APR-1992 (first entry)
XX
DE Sequence of synthetic gene encoding human calcitonin analogue.
XX
KW Paget's disease; therapy; hypercalcaemia; synthetic gene;
KW calcitonin analogue; ds.
XX
OS Homo sapiens.
XX
PN WO8304028-A.
XX
PD 24-NOV-1983.
XX
PF 15-APR-1983; 83WO-0901772.
XX
PR 06-MAY-1982; 82US-0375499.
XX
PA (MOLE-) APPL MOLECULAR GENE.
PA (AMGE-) AMGEN INC.
XX
PI Lau E, Suggs S;
XX
DR WPI; 1983-833187/48.
XX

(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and anti-inflammatory activity. (I), nucleic acids encoding (I), ABM66044-ABN71526 and antibodies that bind (I) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes. Nucleic acids encoding (I) are used to detect Streptococcus in a biological sample. (I) is used to determine whether a compound binds to (I). A composition comprising (I) or a nucleic acid encoding (I), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (I) may be used to recombinantly produce (I) and may be used in gene therapy. Antibodies to (I) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins.

SQ Sequence 93 BP; 33 A; 17 C; 18 G; 25 T; 0 other;
 Query Match 0.9%; Score 24.6; DB 24; Length 93;
 Best Local Similarity 59.2%; Pred. No. 4e+03;
 Matches 42; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy	1620	TGAGGGCAACAACACTATGATACACCTGAGCTGGGACTTTTCCAGCTCTCTCCAGCG	1679
Db	5	TGACAGTAATTGGGAACCATTCTAGAGGTGAAATATATAGTGCTGCCTATTTCCAGTT	64

Qy 1680 ATTCATCAGGA 1690
Db 65 ATTCATCAGAA 75

RESULT 13
AAT23132
ID AAT23132 standard; cDNA to mRNA; 95 BP.

AA AAT23132;
AC

DT 23-AUG-1996 (first entry)

Human gene signature HUMGS04866.

Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function: ss.

OS Homo sapiens.

PN W09514772-A1.

01-JUN-1995.

11-NOV-1994; 94WO-JP01916.

PR 12-NOV-1993; 93JP-0355504.

PA (MATS/) MATSUBARA K.

XX
(UKUB/) UKUBU K.:

FI
XX

XX
WFL; 1993-2009T/27.

PT for diagnosis of abnormal cell function, by preparing cDNA that reflects relative abundance of corresp. mRNA in specific human tissues

PS Claim 1: Page 1296: 2245pp: Japanese.

A single-stranded DNA (or its complementary strand or the corresponding double-stranded DNA) which comprises one of the 7037 "GS" sequences given in AAT19001-T28837 and which is able to hybridise to part of human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)

CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

Sequence 95 BP; 55 A; 10 C; 20 G; 10 T; 0 other;
SQ

Query Match 0.98; Score 24.6; DB 16; Length 95;

Best Local Similarity 39.2%, Fred. NO: 4.1e+03,
Matches 42; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

[illegible]

Qy 2509 GGAGAAGGTGA 2519
 |||
Db 84 TGAAGAAGGTAA 94

RESULT 14
ABA48667
ID ABA48667 standard; DNA; 90 BP.

XX ABA48667;
AC

01-FEB-2002 (first entry)

DE Human breast cell single exon nucleic acid probe #7362.

AA Human; microarray; single exon probe; gene expression; breast.
KW
KW disease; cancer; ss.

OS Homo sapiens.

AX WO200157271-A2.
PN

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00662.

PR 04-FEB-2000; 2000US-0180312.

PR 30-JUN-2000; 2000US-0608408.

PR 21-SEP-2000; 2000US-0234687.

FK 27 SEF 2000; 2000GB-V230353.
PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-496933/54.

New spatially-addressable set of single exon nucleic acid probes, useful for measuring gene expression in sample derived from human breast, comprises number of single exon nucleic acid probes -

PS Claim 4; SEQ ID NO 7362; 327pp + sequence listing; English.

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human breast and BT 474 cells. The method involves contacting the probes with a collection of detectably labelled nucleic acids derived from mRNA of human breast, and then measuring the label

CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 90 BP; 31 A; 24 C; 17 G; 18 T; 0 other;

Query Match 0.9%; Score 24.4; DB 22; Length 90;
Best Local Similarity 56.1%; Pred. No. 4.5e+03;
Matches 46; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 1966 TCTCACAGACCTTCTGCCACTGGGACATGACAATCACGTGCAGCTCAAGTGGAGTGTG 2025
DB 9 TCCTACCCAAAGCTTCTCTCATCTGTGAACAGACAGGAGGCGCTCGTAAATAGGTACATG 68
QY 2026 TTGACAGCAAGAGGGAGCCCA 2047
DB 69 TTGACAGCAACACCTACTCA 90

RESULT 15
ABA66579
ID ABA66579 standard; DNA; 90 BP.
XX
AC ABA66579;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #14884.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
PS Claim 4; SEQ ID NO 14884; 639pp + sequence listing; English.
XX

XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 90 BP; 31 A; 24 C; 17 G; 18 T; 0 other;
Query Match 0.9%; Score 24.4; DB 22; Length 90;
Best Local Similarity 56.1%; Pred. No. 4.5e+03;
Matches 46; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 1966 TCTCACAGACCTTCTGCCACTGGGACATGACAATCACGTGCAGCTCAAGTGGAGTGTG 2025
DB 9 TCCTACCCAAAGCTTCTCTCATCTGTGAACAGACAGGAGGCGCTCGTAAATAGGTACATG 68
QY 2026 TTGACAGCAAGAGGGAGCCCA 2047
DB 69 TTGACAGCAACACCTACTCA 90
Search completed: November 20, 2002, 00:05:47
Job time : 390 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	22.6	0.8	97	1	US-08-479-783A-76	Sequence 76, Appl
2	22.6	0.8	97	1	US-08-479-723-76	Sequence 76, Appl
3	22.6	0.8	97	1	US-08-618-693-76	Sequence 76, Appl
4	22.6	0.8	97	4	US-08-973-124-165	Sequence 165, App
5	22.6	0.8	97	4	US-08-991-743C-76	Sequence 76, Appl
6	22.6	0.8	97	5	PCF-US96-08014-165	Sequence 165, App
7	22.4	0.8	40	1	US-08-209-951-40	Sequence 1, Appl
8	22.4	0.8	40	2	US-08-593-205-1	Sequence 1, Appl
9	22.4	0.8	60	2	US-07-662-764D-34	Sequence 34, Appl
c 10	22.4	0.8	60	2	US-07-662-764D-34	Sequence 34, Appl
11	22.4	0.8	65	2	US-08-418-848A-70	Sequence 70, Appl
c 12	22.4	0.8	75	2	US-08-835-099A-8	Sequence 8, Appl
c 13	22.4	0.8	75	3	US-09-157-345-8	Sequence 8, Appl
14	22.4	0.8	79	1	US-08-446-908-8	Sequence 8, Appl
15	22.4	0.8	79	1	US-08-231-205A-8	Sequence 8, Appl
16	22.4	0.8	79	2	US-08-871-161-8	Sequence 8, Appl
17	22.4	0.8	93	4	US-09-511-625B-40	Sequence 40, Appl
18	22.4	0.8	96	1	US-08-479-783A-67	Sequence 67, Appl
19	22.4	0.8	96	1	US-08-479-725-67	Sequence 67, Appl
20	22.4	0.8	96	1	US-08-618-693-67	Sequence 67, Appl
21	22.4	0.8	96	4	US-08-973-124-156	Sequence 156, App
22	22.4	0.8	96	4	US-08-991-743C-67	Sequence 67, Appl
23	22.4	0.8	96	5	PCF-US96-08014-156	Sequence 156, App
24	22.4	0.8	97	4	US-08-952-793-272	Sequence 272, App
25	22.4	0.8	97	5	PCF-US96-09455A-272	Sequence 272, App
26	22.2	0.8	71	4	US-09-506-729-19	Sequence 19, Appl
27	22.2	0.8	78	3	US-08-718-904-95	Sequence 95, Appl

[illegible]


```
RESULT 8
US-08-593-205-1
; Sequence 1, Application US/08593205
; Patent No. 5866421
; GENERAL INFORMATION:
; APPLICANT: Calgene, Inc.
; TITLE OF INVENTION: ENHANCED EXPRESSION IN A PLANT
; TITLE OF INVENTION: PLASTID
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
; COMPUTER: IBM PC
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: Microsoft word for Windows 7.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/593,205
; FILING DATE: January 29, 1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02901
; FILING DATE: March 10, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 105-1 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (530) 753-6313
; TELEFAX: (530) 753-1510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: synthetic oligonucleotide
; US-08-593-205-1

Query Match 0.8%; Score 22.4; DB 2; Length 40;
Best Local Similarity 72.5%; Pred. No. 2e+03;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 481 CCCGATTCCCTGAAATAATCCCAACAGCCTTGAATGCA 520
||||| | || || |||| |||| | |||||
Db 1 CCCGGATCCATGGATAACAATCCGAACATCAATGAATGCA 40

RESULT 9
US-07-662-764D-34
; Sequence 34, Application US/07662764D
; Patent No. 5866363
; GENERAL INFORMATION:
; APPLICANT: Piezenik, George
; TITLE OF INVENTION: METHOD AND MEANS FOR SORTING AND
; TITLE OF INVENTION: IDENTIFYING BIOLOGICAL INFORMATION
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ &
; ADDRESSEE: MENTLIK
; STREET: 600 South, Avenue West
; CITY: Westfield
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/662,764D
; FILING DATE: 28-FEB-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/201,358
; FILING DATE: 26-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/770,390
; FILING DATE: 28-AUG-1985
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/662,764D
; FILING DATE: 28-FEB-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/201,358
; FILING DATE: 26-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/770,390
; FILING DATE: 28-AUG-1985
; ATTORNEY/AGENT INFORMATION:
; NAME: Foley, Shawn P.
; REGISTRATION NUMBER: 33,071
; REFERENCE/DOCKET NUMBER: ICTECH/0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-654-5000
; TELEFAX: 908-654-7866
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 41..58
; US-07-662-764D-34

Query Match 0.8%; Score 22.4; DB 2; Length 60;
Best Local Similarity 62.5%; Pred. No. 2.6e+03;
Matches 35; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1509 GAAGCACCAGAGAACAAAGTGTTCATGAGGAAGTTCAAGATCGGTACAGCAACA 1564
||||| | || || |||| |||| | |||||
Db 5 GATGTACTTTGCTTGAAGATCCTCATGAGGATCTTCAAGACAAAGTACATCAAGA 60

RESULT 10
US-07-662-764D-34/c
; Sequence 34, Application US/07662764D'
; Patent No. 5866363
; GENERAL INFORMATION:
; APPLICANT: Piezenik, George
; TITLE OF INVENTION: METHOD AND MEANS FOR SORTING AND
; TITLE OF INVENTION: IDENTIFYING BIOLOGICAL INFORMATION
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ &
; ADDRESSEE: MENTLIK
; STREET: 600 South, Avenue West
; CITY: Westfield
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/662,764D
; FILING DATE: 28-FEB-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/201,358
; FILING DATE: 26-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/770,390
; FILING DATE: 28-AUG-1985
```



```

RESULT 14
US-08-446-908-8
Sequence 8, Application US/08446908
Patent No. 5705149
GENERAL INFORMATION:
APPLICANT: Namen, Anthony E.
APPLICANT: Goodwin, Raymond G.
APPLICANT: Lupton, Stephen D.
APPLICANT: Mochizuki, Diane Y.
TITLE OF INVENTION: Interleukin
TITLE OF INVENTION: Therewith
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

```

RESULT 15
US-08-231-205A-8
; Sequence 8, Application US/08231205A
; Patent No. 5714585
; GENERAL INFORMATION:
; APPLICANT: Namen, Anthony E.
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Lupton, Stephen D.
; APPLICANT: Mochizuki, Diane Y.
; TITLE OF INVENTION: Interleukin-7 and Antibodies Reactive
; TITLE OF INVENTION: Therewith
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple 7.1
; SOFTWARE: Microsoft Word, Version 5.1a

```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,205A
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/957,649
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,438
; FILING DATE: 13-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/255,209
; FILING DATE: 07-OCT-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/113,566
; FILING DATE: 26-OCT-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2104-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 19..60
; US-08-231-205A-8

Query Match      0.8%; Score 22.4; DB 1; Length 79;
Best Local Similarity 66.7%; Pred. No. 3.le+03;
Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 228 TTGTGGAGGACAGACTGCAAGTATGACTACGTGGAAGTCTTCGATGG 275
      ||||| | ||||| |||| | || | | | |||| |
Db 6 TTGTGATAAAGAGACTACAGGACCGCATGCAAGGAATCGAAGG 53

Search completed: November 20, 2002, 00:07:23
Job time : 80 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2002, 20:45:11 ; Search time 4805 Seconds
(without alignments)
16789.381 Million cell updates/sec

Title: US-09-296-264-33
Perfect score: 2772
Sequence: 1 atggagaggggctgcgct.....gtactattcggagggcatga 2772

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2772	100.0	2772	9	AF018956 Homo sapi
2	2760.8	99.6	2772	9	AF016050 Homo sapi
3	2741.4	98.9	2775	6	AX133541 Sequence
4	2536.6	91.5	2568	6	AX133545 Sequence
5	2338	84.3	5421	9	HSMB04921
6	2258	81.5	2808	6	AX147495 Sequence
7	2210.4	79.7	3471	10	AF016296 Rattus no
8	2202.4	79.5	3652	6	AX305466 Sequence
9	2202.4	79.5	3652	10	MUSMONE
10	2196	79.2	2766	10	AF018957
11	2026.2	73.1	2679	6	ES9778 Soluble neu
12	1920.8	69.3	2232	9	BC007533 Homo sapi
13	1920.8	69.3	2241	6	AX147494 Sequence
14	1919.2	69.2	1972	9	AF145712 Homo sapi
15	1862.2	67.2	3750	9	AF280547 Homo sapi
16	1760.2	63.5	2014	9	BC007737 Homo sapi
17	1676	60.5	3854	5	CHKVII
18	1610.2	58.1	1922	6	AX147496 Sequence
19	1437.2	51.8	3800	5	XELA5P
20	1139.8	41.1	1550	6	AX147488 Sequence
21	1135.6	41.0	3174	5	AY064213 Danio rer
22	982.4	35.4	1550	6	AX147489 Sequence
23	596.2	21.5	4733	10	AF022861 Mus muscu
24	583	21.0	4718	10	AF022854 Mus muscu
25	570.6	20.6	2730	9	AF022859 Homo sapi
26	568.8	20.5	2760	5	AF417235 Gallus ga
27	566.2	20.4	4769	10	AF022855 Mus muscu
28	565.4	20.4	2781	10	AF483506 Mus muscu
29	565.4	20.4	2781	10	AF483507 Mus muscu
30	563.6	20.3	3371	10	AF016297 Rattus no
31	561.8	20.3	2781	9	AF022860 Homo sapi
32	552.4	19.9	4784	10	AF022856 Mus muscu
33	549.6	19.8	2796	9	AF016098 Homo sapi
34	548	19.8	6690	9	HSMB04919
35	537.8	19.4	2670	6	ES9779 Soluble neu
36	537.8	19.4	4765	10	AF022857 Mus muscu
37	537.8	19.4	4780	10	AF022858 Mus muscu
38	537	19.4	3357	9	AF280544 Homo sapi
39	537	19.4	3372	9	AF280545 Homo sapi
40	531	19.2	3186	4	AF395335 Ovis arie
41	529.8	19.1	2811	5	AF417236 Gallus ga
42	490	17.7	1810	9	AF280546 Homo sapi
43	333.4	12.0	178183	9	HSB342D11 Human DNA
C 44	333.4	12.0	179684	2	AL591369 Homo sapi
C 45	274	9.9	861	9	AF280548S4 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AF018956
DEFINITION Homo sapiens neuroligin mRNA, complete cds.
ACCESSION AF018956
VERSION AF018956.1 GI:2407640
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2772)
AUTHORS He, Z. and Tessier-Lavigne, M.
TITLE Neuroligin is a receptor for the axonal chemorepellent Semaphorin III

PRI 18-SEP-1997

JOURNAL Cell 90 (4), 739-751 (1997)
MEDLINE 97433084
PUBMED 9288753
REFERENCE 2 (bases 1 to 2772)
AUTHORS He,Z. and Tessier-Lavigne,M.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1997) Howard Hughes Medical Institute, University
of California, San Francisco, 513 Parnassus Avenue, HSE-201, San
Francisco, CA 94143, USA
FEATURES Location/Qualifiers
source 1..2772
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 1..2772
/product="neuropilin"
/protein_id="AAC51759.1"
/db_xref="GI:2407841"
/translation="MERGLPLLCALVLAIPAGAFRNDCEGDTIKIESPGYLTPGY
PHSYHPSKCEWLIOADPYQIRIMINFNPHFLEDRDCKYDVEVFDGENENGHPRGK
FCGKIAPPVSVSSGPELFKVFYDSEYELHFEFSDLEPDSNPPGMCRCYDRLEINDPFP
PGFPEKYNSELECTIVFAPKMSSEIILEFESFDEPDSNPPGMCRCYDRLEINDPFP
DVGPHGYCCQKTPGRLRSSGILLSMVFYDSEYELHFEFSDLEPDSNPPGMCRCYDRLEINDPFP
EALHMESEIISDQITASSQYSTNWSAERSRLNYPENGWTPEDSIFRWIQVDLGLLR
FVAVGTGSAISKETKKYKYKTYKIDVSSNGEDWITKEGKNPVLFOGNTPTDVVY
AVFPPLTIRFVRIRKATWETGISMRFEVYCKITDPCSGMLGWSGLISDSQITSS
NOGDRNMPENIRLVTSRSGWALPPAPHSYINELQIDLBEKIVRGIIIOGKHREN
KVFMKFGIYVSSNGSDMKMTDDSKRAKSFEGNNYDTPELRFPALSTRFIRIYP
ERATHGKGLBELLGCEVAPTAGPTTPNGNLVDECDQDQANCHSGTGDQFOLTGGT
TVLATEKPTVIDSTIQSEFPYTFNCFEGWSHKTFCHEHDNHNVLKWSVLTSKGP
IODHFGDNFIYSADENOKGVARLVSPVYVYSONSAHCFMTFWHMSGSHVGTLRVLL
RYQKPEEDQVLVWMAIQHGDHWKEGRVLLHLSKLYQVIFEIGLGRNLGAGIADVDI
SINHHISODCAKPADLDKNPEIKIDETGTPGYEGEGEDKNISRPKNVLTLEP
ILITIIAMSAKVLGLGCVGVLYLCACWHNGMSERNLSALENYNPELVGVKLLKDKL
NTQSPYSER"
BASE COUNT 772 a -664 c 702 g 634 t
ORIGIN
Query Match 100.0%; Score 2772; DB 9; Length 2772;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2772; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGAGGGGGTGGCGCTCTCTGCGCGGTGCTCGCCCTCGTCTCGCCCGCGCGGC 60
DB 1 ATGGAGAGGGGGTGGCGCTCTCTGCGCGGTGCTCGCCCTCGTCTCGCCCGCGCGGC 60
QY 61 GCTTTTCGCAACGATGAATGTGGCGATCTATATAAATTTGAAGCCCGGGTACCTTACA 120
DB 61 GCTTTTCGCAACGATGAATGTGGCGATCTATATAAATTTGAAGCCCGGGTACCTTACA 120
QY 121 TCTCTGGTTATCTCTTATCACCAAGTGAAGAAATGGAATGGCTGATTCAGGCT 180
DB 121 TCTCTGGTTATCTCTTATCACCAAGTGAAGAAATGGAATGGCTGATTCAGGCT 180
QY 181 CCGGACCCATACAGAGAATTATGATCACTTCAACCCCTCACTTCGATTTGGAGGACAGA 240
DB 181 CCGGACCCATACAGAGAATTATGATCACTTCAACCCCTCACTTCGATTTGGAGGACAGA 240
QY 241 GACTGCAAGTATGACTTACCTGGAAGTCTTCGATGGAGAAATGAAATGGACATTTTAGG 300
DB 241 GACTGCAAGTATGACTTACCTGGAAGTCTTCGATGGAGAAATGAAATGGACATTTTAGG 300
QY 301 GGAAAGTTCCTGGAAGATAGCCCTCTCTCTGTTGTCTTCAGGGGCATTTCTTTT 360
DB 301 GGAAAGTTCCTGGAAGATAGCCCTCTCTCTGTTGTCTTCAGGGGCATTTCTTTT 360
QY 361 ATCAAAATTTGCTCTGACTACGAAACACATGTTGCGAGGATTTTCCATACGTTATGAAAT 420
DB 361 ATCAAAATTTGCTCTGACTACGAAACACATGTTGCGAGGATTTTCCATACGTTATGAAAT 420
QY 421 TTCAAGAGAGGTCCTGAATGTTTCCAGAGACTACAAACACCTAGTGGAGTGAATAGTCC 480
DB 421 TTCAAGAGAGGTCCTGAATGTTTCCAGAGACTACAAACACCTAGTGGAGTGAATAGTCC 480

QY 481 CCCGGATTCCTCGAAAAATATCCCAACAGCCTTGAATGCACCTTATATGTCTTTGGGCCA 540
DB 481 CCCGGATTCCTCGAAAAATATCCCAACAGCCTTGAATGCACCTTATATGTCTTTGGGCCA 540
QY 541 AGATCTCAGAGATTTATCTGGAAATTTGAAGCTTTGACCTGGAGCCTGACTCAAAATCCT 600
DB 541 AGATCTCAGAGATTTATCTGGAAATTTGAAGCTTTGACCTGGAGCCTGACTCAAAATCCT 600
QY 601 CCAGGGGGGATGTTCTGTCGCTACGACCGGCTAGAAATCTGGATCGATCCCTGATGTT 660
DB 601 CCAGGGGGGATGTTCTGTCGCTACGACCGGCTAGAAATCTGGATCGATCCCTGATGTT 660
QY 661 GGCCCTCACATTTGGGCGTTACTGTGACAGAAAAACACAGGTGGAATCCGATCCTCATCG 720
DB 661 GGCCCTCACATTTGGGCGTTACTGTGACAGAAAAACACAGGTGGAATCCGATCCTCATCG 720
QY 721 GGCAATTCCTCATGTTTATACCCGACGCGCATAGCAAAAGAAGTTCCTCAGCA 780
DB 721 GGCAATTCCTCATGTTTATACCCGACGCGCATAGCAAAAGAAGTTCCTCAGCA 780
QY 781 AACTACAGTGTCTTGACAGACGAGTGTCTCAGAGATTTTCAAAATGTATGGAAGCTCTGGGC 840
DB 781 AACTACAGTGTCTTGACAGACGAGTGTCTCAGAGATTTTCAAAATGTATGGAAGCTCTGGGC 840
QY 841 ATGGAATCAGGAGAAATTCATTTCTGACACGATCACAGCTTCTTCCCAGTATAGCAACAAC 900
DB 841 ATGGAATCAGGAGAAATTCATTTCTGACACGATCACAGCTTCTTCCCAGTATAGCAACAAC 900
QY 901 TGGTCTGACAGAGCGTCCCGCTGAACTACCTGAGAAATGGGTGGACTCCCGGAGAGGAT 960
DB 901 TGGTCTGACAGAGCGTCCCGCTGAACTACCTGAGAAATGGGTGGACTCCCGGAGAGGAT 960
QY 961 TCCTACGAGAGTGGATACAGTAGACTTGGCGCTTCTCGCTTCTGCTGAGCTGCTCGG 1020
DB 961 TCCTACGAGAGTGGATACAGTAGACTTGGCGCTTCTCGCTTCTGCTGAGCTGCTCGG 1020
QY 1021 ACACAGGGCGCCATTTCAAAAAGAAACCAAGAGAAATATTTATGTCAGAGCTTACAAGATC 1080
DB 1021 ACACAGGGCGCCATTTCAAAAAGAAACCAAGAGAAATATTTATGTCAGAGCTTACAAGATC 1080
QY 1081 GACGTTAGTCTCAACGGGGAAGACTGGATCACCATATAAAGAGAAACAAACCTGTTCTC 1140
DB 1081 GACGTTAGTCTCAACGGGGAAGACTGGATCACCATATAAAGAGAAACAAACCTGTTCTC 1140
QY 1141 TTTACAGGAAACACCAACCCACAGATGTTGGTTGCGAGTATTTCCCAAAACCACTGATA 1200
DB 1141 TTTACAGGAAACACCAACCCACAGATGTTGGTTGCGAGTATTTCCCAAAACCACTGATA 1200
QY 1201 ACTCGATTTGTCCGAATCAAGCTTGCACCTTGGAAACTGGCATATCTATGAGATTGAA 1260
DB 1201 ACTCGATTTGTCCGAATCAAGCTTGCACCTTGGAAACTGGCATATCTATGAGATTGAA 1260
QY 1261 GTATACGGTTGCAAGATACAGATATTCCTGCTCTGGAATGTTGGGTATGGTGTGGA 1320
DB 1261 GTATACGGTTGCAAGATACAGATATTCCTGCTCTGGAATGTTGGGTATGGTGTGGA 1320
QY 1321 CTTATTTCTGACTCCAGATCACATCATCAACAGGAGACAGAACTGGATGCTGAA 1380
DB 1321 CTTATTTCTGACTCCAGATCACATCATCAACAGGAGACAGAACTGGATGCTGAA 1380
QY 1381 AACATCCGCTTGGTACCAAGTCTGCTGCTGGGCACTTCCACCCGACCTCATTTCTTAC 1440
DB 1381 AACATCCGCTTGGTACCAAGTCTGCTGCTGGGCACTTCCACCCGACCTCATTTCTTAC 1440
QY 1441 ATCAATAGTGGTCCAAATAGACCTGGGGGAGGAGAGATCGTGGAGGGGATCATCAT 1500
DB 1441 ATCAATAGTGGTCCAAATAGACCTGGGGGAGGAGAGATCGTGGAGGGGATCATCAT 1500
QY 1501 CAGGCTGGGAACACCGGAGACAGAGTGTTCATCAGGAAGTTCAGATCGGTFACAGC 1560
DB 1501 CAGGCTGGGAACACCGGAGACAGAGTGTTCATCAGGAAGTTCAGATCGGTFACAGC 1560

QY 1561 AACAAAGCGCTCGGACTGGAAGATGATCATGGATGACAGCAACGCAAGCGGAAGTCTTTT 1620
Db 1561 AACAAAGCGCTCGGACTGGAAGATGATCATGGATGACAGCAACGCAAGCGGAAGTCTTTT 1620
QY 1621 GAGGCACACAACTATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCAGCGGA 1680
Db 1621 GAGGCACACAACTATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCAGCGGA 1680
QY 1681 TTCATCAGGATCTACCCCGAGAGAGCCACTCATGCGGACTGGGGCTCAGAAATGAGCTG 1740
Db 1681 TTCATCAGGATCTACCCCGAGAGAGCCACTCATGCGGACTGGGGCTCAGAAATGAGCTG 1740
QY 1741 CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCGACCTCCCAACGGAACCTTGGTG 1800
Db 1741 CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCGACCTCCCAACGGAACCTTGGTG 1800
QY 1801 GATGAATGTGATGACGACGAGCCCAACTGCCACAGTGAACAGGTGATCTCCAGCTC 1860
Db 1801 GATGAATGTGATGACGACGAGCCCAACTGCCACAGTGAACAGGTGATCTCCAGCTC 1860
QY 1861 ACAGGTGGCACCATCTGTCTGCCACAGAAAAGCCCGCTCATAGACGACCAACATAAA 1920
Db 1861 ACAGGTGGCACCATCTGTCTGCCACAGAAAAGCCCGCTCATAGACGACCAACATAAA 1920
QY 1921 TCAGAGTTTCCAACATATGGTTTAACTGTGAATTTGGCTGGGCTCTCACAAAGACCTTC 1980
Db 1921 TCAGAGTTTCCAACATATGGTTTAACTGTGAATTTGGCTGGGCTCTCACAAAGACCTTC 1980
QY 1981 TGCCACTGGACATGACAACTACCTGACGTCAAGTGGAGTGTGTTGACCAAGCAAGAG 2040
Db 1981 TGCCACTGGACATGACAACTACCTGACGTCAAGTGGAGTGTGTTGACCAAGCAAGAG 2040
QY 2041 GGACCCATTCAGGATCACACAGGAGATGGCAACTTCATCTATTCCCAAGCTGACCAAAAT 2100
Db 2041 GGACCCATTCAGGATCACACAGGAGATGGCAACTTCATCTATTCCCAAGCTGACCAAAAT 2100
QY 2101 CAGAAAGGCAAGTGGCTCGCTGGTGGACCCCTGTGGTTTATTCACCAAGCTCTGCCAC 2160
Db 2101 CAGAAAGGCAAGTGGCTCGCTGGTGGACCCCTGTGGTTTATTCACCAAGCTCTGCCAC 2160
QY 2161 TGCATGACCTTCTGGTATACATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2220
Db 2161 TGCATGACCTTCTGGTATACATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2220
QY 2221 CGCTACCAAGAGGAGTACCATGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2280
Db 2221 CGCTACCAAGAGGAGTACCATGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2280
QY 2281 GACCACTGGAAGGAGGCGTGTCTTGTCCACAACTCTCTGAACTTTATCAGGTGAT 2340
Db 2281 GACCACTGGAAGGAGGCGTGTCTTGTCCACAACTCTCTGAACTTTATCAGGTGAT 2340
QY 2341 TTCAGAGGCGAAATCGGAAAGGAAACCTTGGTGGATTTGCTGTGGATGACATTAGTATT 2400
Db 2341 TTCAGAGGCGAAATCGGAAAGGAAACCTTGGTGGATTTGCTGTGGATGACATTAGTATT 2400
QY 2401 AATAACCACTTTCACAAAGAGATTGTGCAAAACAGCAGACCTGGATGATGAAAGAACCCA 2460
Db 2401 AATAACCACTTTCACAAAGAGATTGTGCAAAACAGCAGACCTGGATGATGAAAGAACCCA 2460
QY 2461 GAAATTAATTTGATGAAACAGGAGCAGCGCAGGATACGAAAGGTGAAGAGAGAGGTGAC 2520
Db 2461 GAAATTAATTTGATGAAACAGGAGCAGCGCAGGATACGAAAGGTGAAGAGAGAGGTGAC 2520
QY 2521 AGAACATCTCCAGGAGCCAGGCAATGTTGAAGACCTTAGAACCCATCCTCATCACC 2580
Db 2521 AGAACATCTCCAGGAGCCAGGCAATGTTGAAGACCTTAGAACCCATCCTCATCACC 2580
QY 2581 ATCATAGGCATGAGCGCCCTGGGGTCTCTCTGGGGTCTCTGTGGGTCTGTGGTGTGAC 2640
Db 2581 ATCATAGGCATGAGCGCCCTGGGGTCTCTCTGGGGTCTCTGTGGGTCTGTGGTGTGAC 2640
QY 2641 TGTGCTGTGGCATAAATGGGATGTGAGAAAGAAACTTGTCTGCCCTGGAGAACTATAAC 2700

Db 2641 TGTGCTGTGGCATAAATGGGATGTGAGAAAGAAACTTGTCTGCCCTGGAGAACTATAAC 2700
QY 2701 TTTGAACTTGTGGATGGTGTGAAGTTGAAAAAGACAACTCAATACACAGAGTACTTAT 2760
Db 2701 TTTGAACTTGTGGATGGTGTGAAGTTGAAAAAGACAACTCAATACACAGAGTACTTAT 2760
QY 2761 TCGGAGGCATGA 2772
Db 2761 TCGGAGGCATGA 2772
RESULT 2
AF016050
LOCUS Homo sapiens vascular endothelial cell growth factor 165
DEFINITION AF016050
ACCESSION AF016050
VERSION AF016050.1 GI:2978559
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2772)
AUTHORS Soker,S., Takashima,S., Miao,H.Q., Neufeld,G. and Klagsbrun,M.
TITLE Neuropilin-1 is expressed by endothelial and tumor cells as an
isoform-specific receptor for vascular endothelial growth factor
JOURNAL Cell 92 (6), 735-745 (1998)
MEDLINE 98188099
PUBMED 9529250
REFERENCE 2 (bases 1 to 2772)
AUTHORS Soker,S., Takashima,S. and Klagsbrun,M.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-1997) Surgical Research, Children's Hospital, 320
Longwood Ave., Boston, MA 02115, USA
FEATURES
source
1. .2772
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="MDA-MB-231"
/cell_type="adenocarcinoma"
/tissue_type="breast"
1. .2772
/gene="VEGF156"
1. .2772
/gene="VEGF156"
/codon_start=1
/product="vascular endothelial cell growth factor 165
receptor/neuropilin"
/protein_id="AAC12921.1"
/db_xref="GI:2978560"
/translation="MERGLPLLCALVLAALPACAFNRDKCGDTIKIESPCYLTSPGY
PHSTHPSEKCEWLIQADPYQIRIMINPHEDLEDRCKDYIEVFGEENGNGHFRGK
FCGFIAPPPVNSGGFLFIKEVSDYETHAGFSIRYIFKRGPCCSNYTPPSGVKIS
PGFPEKYPNSLECTYIVFAPKMSIILEFESFLEPDSNPPGFCFDRLEIWDGFP
DVGHIGRYCGKTPGRISSSGILSMVFYDLSAIKAGFSANTSVLQSSYSEDFKCM
EALGMSETHSDOITASSOYSTNWSAERSRLNYPENGWTPGSDSPYQVLDGLLR
FVAVTCGSAISKETKKYKVKYTKIDVSSNGEDWITIKENKPVLEQGNPTDVVY
AVFPKLITFRVIRKATWETGISMREFYVCKITIDYPCSNLGMVSGLISDSQTS
NQDRNMPENIRLVRTSRGALPAPHSYINELQIDLGEEKIVRGIIQGGKHREN
KRVFRKFGIGYSNNGSDCKMIMDDSKRAKSFEGNNYDTPELRTFPALSTRIYIP
ERATHGGLGRMELLGCEVEAPTAGTTPGNLVNDECDQOANCHSGTGDQFLTGGT
TVLATEKPTVIDSTIQSEFPYGFCEFGMSHKTFCHEHDHNVOLKWSVLTSGTGP
IQDHTGDNFIYSOADENQKVARLSPVYVYSONSAHCFTFWYHMSGSHVGTLRVLL
RYQKPEETHQLVWMAHQGDHWRVLLHKSUKLVQVIFEGEIGKNGJGGLIADDD
SINNHISQEDCAPDLDKNPKIKIDETGTPGEGEGEGDKNISRRKPGVNLKLDLP
ILITIIAMSAKGLVLLGVCVGVLYCACWHNGMSERNLSALENYFELVDGVKLLKKDL
NTQSTYSEA"
BASE COUNT 771 a 665 c 702 g 634 t
ORIGIN
Query Match 99.6%; Score 2760.8; DB 9; Length 2772;

Best Local Similarity 99.7%; Pred. No. 0;			
Matches 2765; Conservative 0; Mismatches 7; Indels 0; Gaps 0;			
QY	1	ATCGAGAGGGGGTGGCGCTCTCGCGCGTCTCGCCCTCGCTCGCTCGCGCGCGCGG	60
Db	1	ATCGAGAGGGGGTGGCGCTCTCGCGCGTCTCGCCCTCGCTCGCGCGCGCGG	60
QY	61	GCCTTTTCGCAACGATGAATGTGCGTACTATATAAAATTGAAAGCCCGCGGTACCTTACA	120
Db	61	GCCTTTTCGCAACGATGAATGTGCGTACTATATAAAATTGAAAGCCCGCGGTACCTTACA	120
QY	121	TCTCCTGGTTATCCTCATTTTATACCCCAAGTGAAATGCGAATGGCTGATTCAGGCT	180
Db	121	TCTCCTGGTTATCCTCATTTTATACCCCAAGTGAAATGCGAATGGCTGATTCAGGCT	180
QY	181	CCGGACCCATACAGAGAAATATGATCACTTCAACCCCTCACCTTCGATTTGGAGGACAGA	240
Db	181	CCGGACCCATACAGAGAAATATGATCACTTCAACCCCTCACCTTCGATTTGGAGGACAGA	240
QY	241	GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAATGAAATGGACATTTTAGG	300
Db	241	GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAATGAAATGGACATTTTAGG	300
QY	301	GGAAAGTCTGTGGAAAGATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	360
Db	301	GGAAAGTCTGTGGAAAGATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	360
QY	361	ATCAAAATTTGCTCTGACTACGAAACACATGGTGCAGGATTTTCCATAGCTTATGAAAT	420
Db	361	ATCAAAATTTGCTCTGACTACGAAACACATGGTGCAGGATTTTCCATAGCTTATGAAAT	420
QY	421	TTCAAGAGAGGTCTCTGAAATTTCCAGAACTACACACACCTAGTGGAGTGATAAGTCC	480
Db	421	TTCAAGAGAGGTCTCTGAAATTTCCAGAACTACACACACCTAGTGGAGTGATAAGTCC	480
QY	481	CCCGGATTCCTGAAATATCCACAGCCTTGAATGCATTTATTTGCTTTGGCGCA	540
Db	481	CCCGGATTCCTGAAATATCCACAGCCTTGAATGCATTTATTTGCTTTGGCGCA	540
QY	541	AGATCTCAGAGATATCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAAATCCT	600
Db	541	AGATCTCAGAGATATCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAAATCCT	600
QY	601	CCAGGGGGATGTTCTGCTACGACCGCTAGAAATCTGGGATGGATTTCCCTGATGT	660
Db	601	CCAGGGGGATGTTCTGCTACGACCGCTAGAAATCTGGGATGGATTTCCCTGATGT	660
QY	661	GGCCCTCATTGGGGCTTACTGTGACAGAAACACAGGTCGAATCCGATCCTCATCG	720
Db	661	GGCCCTCATTGGGGCTTACTGTGACAGAAACACAGGTCGAATCCGATCCTCATCG	720
QY	721	GGCATTCTCTCCATGGTTTTTACACCGACAGCGCGATAGCAAAAGAGTTTCTCAGCA	780
Db	721	GGCATTCTCTCCATGGTTTTTACACCGACAGCGCGATAGCAAAAGAGTTTCTCAGCA	780
QY	781	AACTACAGTGTCTGACAGCAGTGTCTCAGAGATTTTCAAAATGTATGGAAGCTCTGGC	840
Db	781	AACTACAGTGTCTGACAGCAGTGTCTCAGAGATTTTCAAAATGTATGGAAGCTCTGGC	840
QY	841	ATGGAATCAGGAGAAATTCATTTGACAGATACAGCTTCTCCAGTATAGCACCAC	900
Db	841	ATGGAATCAGGAGAAATTCATTTGACAGATACAGCTTCTCCAGTATAGCACCAC	900
QY	901	TGGTCTGACAGAGCGCTCCCGCTGAACTACCTGAGAATGGTGGACTCCCGAGAGGAT	960
Db	901	TGGTCTGACAGAGCGCTCCCGCTGAACTACCTGAGAATGGTGGACTCCCGAGAGGAT	960
QY	961	TCTTACCGAGAGTGGATACAGTGTGGCCCTTCTCGCTTTGTCTACGGCTGTGCGG	1020
Db	961	TCTTACCGAGAGTGGATACAGTGTGGCCCTTCTCGCTTTGTCTACGGCTGTGCGG	1020
QY	1021	ACACAGGGCCCATTTCAAAAGAACCAAGAGAAATATATGTCAGAGCTTACAGATC	1080
Db	1021	ACACAGGGCCCATTTCAAAAGAACCAAGAGAAATATATGTCAGAGCTTACAGATC	1080

Db	1021	ACACAGGGCCCATTTCAAAAGAACCAAGAGAAATATATGTCAGAGCTTACAGATC	1080
QY	1081	GAGCTTAGCTCCCAACGGGGAAGACTGGATCACCATAAAAGAAACAAACCTGTTCTC	1140
Db	1081	GAGCTTAGCTCCCAACGGGGAAGACTGGATCACCATAAAAGAAACAAACCTGTTCTC	1140
QY	1141	TTTCAGGGAACACCAACCCACAGATGTTGTGTGTGACATTTCCCCAAACCACTGATA	1200
Db	1141	TTTCAGGGAACACCAACCCACAGATGTTGTGTGTGACATTTCCCCAAACCACTGATA	1200
QY	1201	ACTCGATTTGTCGATCAAGCCTGCAACTGGGAACCTGGCATATCTATGAGATTGAA	1260
Db	1201	ACTCGATTTGTCGATCAAGCCTGCAACTGGGAACCTGGCATATCTATGAGATTGAA	1260
QY	1261	GTATACGGTTGCAAGATACAGATATTCCTTCTCTGGAATGTTGGTATGGTGTGGA	1320
Db	1261	GTATACGGTTGCAAGATACAGATATTCCTTCTCTGGAATGTTGGTATGGTGTGGA	1320
QY	1321	CTTATTTCTGACTCCAGATCACATCATCAACCAAGGAGACAGAACTGGATGCCGTAA	1380
Db	1321	CTTATTTCTGACTCCAGATCACATCATCAACCAAGGAGACAGAACTGGATGCCGTAA	1380
QY	1381	AACATCCGCTGTATACAGTCTGCTGGCTGGGACTTCCACCCGACCTCATTCCTAC	1440
Db	1381	AACATCCGCTGTATACAGTCTGCTGGCTGGGACTTCCACCCGACCTCATTCCTAC	1440
QY	1441	ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAGATCGTGAGGGGCATCATCAT	1500
Db	1441	ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAGATCGTGAGGGGCATCATCAT	1500
QY	1501	CAGGGTGGGAAGCACCGAGAGAACAAAGGTGTTTCATGAGGAAGTTCAAGATCGGATAC	1560
Db	1501	CAGGGTGGGAAGCACCGAGAGAACAAAGGTGTTTCATGAGGAAGTTCAAGATCGGATAC	1560
QY	1561	AACAACGGCTCGGACTGGAAGATGATCATGGATGAGACGAAACGCAAGGGAAGTCTTT	1620
Db	1561	AACAACGGCTCGGACTGGAAGATGATCATGGATGAGACGAAACGCAAGGGAAGTCTTT	1620
QY	1621	GAGGGCAACAACTATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCCACGCGA	1680
Db	1621	GAGGGCAACAACTATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCCACGCGA	1680
QY	1681	TTCATCAGGATCTACCCGAGAGAGCCTCATGTCGGGCTGAGGCTGAGATGGAGCTG	1740
Db	1681	TTCATCAGGATCTACCCGAGAGAGCCTCATGTCGGGCTGAGGCTGAGATGGAGCTG	1740
QY	1741	CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCCACCACTCCCAACGGGAACCTTGG	1800
Db	1741	CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCCACCACTCCCAACGGGAACCTTGG	1800
QY	1801	GATGAATGTGATGACAGCAGGCGCAACTGCGCAGTGGAAACAGGTGATGACTTCCAGCTC	1860
Db	1801	GATGAATGTGATGACAGCAGGCGCAACTGCGCAGTGGAAACAGGTGATGACTTCCAGCTC	1860
QY	1861	ACAGGTGGCAACCACTGTGCTGGCCACAGAAAACCCACGCTCATAGACAGCACCATA	1920
Db	1861	ACAGGTGGCAACCACTGTGCTGGCCACAGAAAACCCACGCTCATAGACAGCACCATA	1920
QY	1921	TGAGAGTTTCCAACTATGTTTAACTGTAATTTGGCTGGGCTCTCACAAGACCTTC	1980
Db	1921	TGAGAGTTTCCAACTATGTTTAACTGTAATTTGGCTGGGCTCTCACAAGACCTTC	1980
QY	1981	TGCACTGGGAACATGACAATCACGTGCGAGCTCAAGTGGAGTGTGTGACCAAGACG	2040
Db	1981	TGCACTGGGAACATGACAATCACGTGCGAGCTCAAGTGGAGTGTGTGACCAAGACG	2040
QY	2041	GGACCCATTCAGGATCACACAGGAGATGGCAACTCATCTATTTCCCAAGCTGACGAAAT	2100
Db	2041	GGACCCATTCAGGATCACACAGGAGATGGCAACTCATCTATTTCCCAAGCTGACGAAAT	2100
QY	2101	CAGAAGGGCAAGTGGCTCGCTGGTGGAGCCCTGTGGTTTATTTCCAGAACTTGGCCAC	2160
Db	2101	CAGAAGGGCAAGTGGCTCGCTGGTGGAGCCCTGTGGTTTATTTCCAGAACTTGGCCAC	2160


```
QY 2161 TCATGACCTTCTGGTATCACATGCTCTGGGTCCTCCAGCTCGGCACACTCAGGCTCAAACTG 2220
Db 2161 TCATGACCTTCTGGTATCACATGCTCTGGGTCCTCCAGCTCGGCACACTCAGGCTCAAACTG 2220
QY 2221 CCTACCAGAACCCAGAGAGTACATCAGCTGGTCTGGATGCCATTTGGACACCAAGGT 2280
Db 2221 CGCTACCAGAACCCAGAGAGTACATCAGCTGGTCTGGATGCCATTTGGACACCAAGGT 2280
QY 2281 GACCACTGGAAGAAAGGGGCTCTTGTCTCCACAAGTCTCTGAACTTTTATCAGGTGATT 2340
Db 2281 GACCACTGGAAGAAAGGGGCTCTTGTCTCCACAAGTCTCTGAACTTTTATCAGGTGATT 2340
QY 2341 TTCGAGGGCGAAATCGGAAAGAAACCTTGTGGATGCTGCTGGATGACATTTAGTATT 2400
Db 2341 TTCGAGGGCGAAATCGGAAAGAAACCTTGTGGATGCTGCTGGATGACATTTAGTATT 2400
QY 2401 AATAACCATTTTACACAGAAGATTGTGCAAAACAGACAGACCTGGATAAAAAGAACCA 2460
Db 2401 AATAACCATTTTACACAGAAGATTGTGCAAAACAGACAGACCTGGATAAAAAGAACCA 2460
QY 2461 GAAATTAATAATTGATGAACAGGGAGCAGCCAGGATAGAAAGTGAAGGAGAGGTGAC 2520
Db 2461 GAAATTAATAATTGATGAACAGGGAGCAGCCAGGATAGAAAGTGAAGGAGAGGTGAC 2520
QY 2521 AAGAACATCTCCAGGAAGCCAGGCAATGTGTTGAAGACCTTAGAACCCATCCTCATCACC 2580
Db 2521 AAGAACATCTCCAGGAAGCCAGGCAATGTGTTGAAGACCTTAGAACCCATCCTCATCACC 2580
QY 2581 ATCATAGCCATCAGCCGCTGGGGTCTCTCTGGGGTCTCTGGGGTCTCTGGGGTCTCTGGT 2640
Db 2581 ATCATAGCCATCAGCGCTGGGGTCTCTCTGGGGTCTCTCTGGGGTCTCTGGGGTCTCTGGT 2640
QY 2641 TGTGCTGTGTCATTAATGGGATGTCAGAAAGAAACTTCTGCTGCCCTGGAGAACTATAAC 2700
Db 2641 TGTGCTGTGTCATTAATGGGATGTCAGAAAGAAACTTCTGCTGCCCTGGAGAACTATAAC 2700
QY 2701 TTTGAACCTTGTGGATGGTGTGAAGTTGAAAAAGACAAACTGAATACACAGAGTACTTAT 2760
Db 2701 TTTGAACCTTGTGGATGGTGTGAAGTTGAAAAAGACAAACTGAATACACAGAGTACTTAT 2760
QY 2761 TCGGAGGCATGA 2772
Db 2761 TCGGAGGCATGA 2772

RESULT 3
AX133541
LOCUS AX133541 2775 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 3 from Patent WO0131346.
ACCESSION AX133541
VERSION AX133541.1 GI:14139691
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2775)
AUTHORS Rosenbaum,J.S., Whitaker,G.B. and Limberg,B.J.
TITLE Identification of novel pro- and anti-angiogenic agents
JOURNAL Patent: WO 0131346-A 3 03-MAY-2001;
THE PROCTER & GAMBLE COMPANY (US)
FEATURES
Location/Qualifiers
source 1. 2775
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 1. 2775
/note="unnamed protein product"
/protein_start=1
/protein_id="CAC38951.1"
/db_xref="GI:14139692"
/translation="MERGLPLLCALVLAIVLAPAGAFNRDKCGDRTIKTESPGYLTSPGY
PHSYHPSEKCEWLIIQAAPPYQRIIMNPHFDLEDRCYDYVEVFDGENENGHFRGK
```

```
FCGIAPPPVSVSGFFLFKFDVYETHGAGFSIRYEIFKRGPECSONYTPSPGVKS
PGFPEKYNSECTYIVPAKMSIILIEFESFDLADPEPDPGCFRDRLEIMDGF
DVGHIGRYCGKTPGRTRSSSGILSMVFTSDATAKEGFSANTSVLOSSEDEKCM
EALGMESEIHSQITASSOYSTNNAEBSRLNYPNGWTPEQDSYREMIQVDGLLR
FVTVAGTGALSKIEKKYKTKIDVSSNGEDWITKEGNKPVLFQGNTPDSDTVV
AVFPRPLITRFRIRKIPATWETGISMRFEVYGGKTDYPCSGMLGVNLSLSDQSIVS
NOGRNMPENIRLVTSRSGWALPAPHSYINELQIDLBEKIVRGIIIOGKHREN
KVWRKFKIGYSNNGSKMIMDDSKRAKSFEGNNYDTPELRTFPALSRFRIRYP
BRATHGGLRLWELLGCBVEAPTAGTTPGNLVCDDDDQDQDQDQDQDQDQDQDQ
TVALATEKPAVISTQSEFPYGFNCFEGWGHSHKTFCHWEHDNHVQLAKWSLTKTGP
IQDHTAGDCNFIYSQADBNQKVARLVSPVYSONSAHCHTFWHMSGSHGLTLRYK
LRYQKPEYDQLVMAHQGHQHWKRGVLLHLSLKYLFQIFEGEKGNGKGIADVD
ISINNHQSQEDCAKPADLDKKNPEIKIDETGTPGYEGEGEDGKNSIKRPNGLATLD
PILITITAMSALGVLGAVCGVLVYCACWHNGMSRNLNLSALENYNFEVLVDGVKLLKDK
LNTOSTYSEA"
BASE COUNT 772 a 563 c 702 g 638 t
ORIGIN
Query Match 98.9%; Score 2741.4; DB 6; Length 2775;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2761; Conservative 0; Mismatches 11; Indels 3; Gaps 1;
QY 1 ATGGAGAGGGGCTGCGCTCCTCTGCGCGCTGCTGCGCCCTGCTCCTCCCGCGCGCGC 60
Db 1 ATGGAGAGGGGCTGCGCTCCTCTGCGCGCTGCTGCGCCCTGCTCCTCCCGCGCGCGC 60
QY 61 GCTTTTCGCAACGATGATGTCGCGATACATAAATAATGAAAGCCCGGGTACCTTACA 120
Db 61 GCTTTTCGCAACGATGATGTCGCGATACATAAATAATGAAAGCCCGGGTACCTTACA 120
QY 121 TCTCTGGTTATCCTCATTTATCACCAGTGAATAATGCGAATGCGTGTATTCAGGCT 180
Db 121 TCTCTGGTTATCCTCATTTATCACCAGTGAATAATGCGAATGCGTGTATTCAGGCT 180
QY 181 CCGGACCCATACAGAGAAATATGATCAACTTCAACCCCTCAGTTCGATTTGGAGGACACA 240
Db 181 CCGGACCCATACAGAGAAATATGATCAACTTCAACCCCTCAGTTCGATTTGGAGGACACA 240
QY 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAATGAAATGGACATTTTAGG 300
Db 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAATGAAATGGACATTTTAGG 300
QY 301 GGAAGTCTGTGGAAAGATAGCCCTCCTCTGTTGTGCTCTCAGGGCCATTTCTTTT 360
Db 301 GGAAGTCTGTGGAAAGATAGCCCTCCTCTGTTGTGCTCTCAGGGCCATTTCTTTT 360
QY 361 ATCAAAATTTGCTCTGACTACGAAACACATGGTGCAGGATTTTCCATAGTTATGAAAT 420
Db 361 ATCAAAATTTGCTCTGACTACGAAACACATGGTGCAGGATTTTCCATAGTTATGAAAT 420
QY 421 TTCAAGAGAGGTCCTGAAATGTTCCAGAACTACACAACTAGTGGAGTATAAGTCC 480
Db 421 TTCAAGAGAGGTCCTGAAATGTTCCAGAACTACACAACTAGTGGAGTATAAGTCC 480
QY 481 CCGGATTCCTTGAAAAATATCCCAACGCTTGAATGCACTATATTTCTTTTGGCCCA 540
Db 481 CCGGATTCCTTGAAAAATATCCCAACGCTTGAATGCACTATATTTCTTTTGGCCCA 540
QY 541 AAGATGTCAGAGTATTCCTGGAATTTGAAAGCTTTGACCTGGAGCTGACTCAATCT 600
Db 541 AAGATGTCAGAGTATTCCTGGAATTTGAAAGCTTTGACCTGGAGCTGACTCAATCT 600
QY 601 CCAGGGGGATGCTCTGCTGCTACGACCGCTAGAAATCTGGAGTGGATTCCTGTAT 660
Db 601 CCAGGGGGATGCTCTGCTGCTACGACCGCTAGAAATCTGGAGTGGATTCCTGTAT 660
QY 661 GGCCCTCAGATTTGGGGTCTTACTGTGGACAGAAAAACACAGGTCGAATCCGATCCTCATCG 720
Db 661 GGCCCTCAGATTTGGGGTCTTACTGTGGACAGAAAAACACAGGTCGAATCCGATCCTCATCG 720
QY 721 GCATTTCTCCTCATGGTTTTTTTACACCGACACCGCGATAGCAAAAGAGGTTTCTCAGCA 780
Db 721 GGCATTTCTCCTCATGGTTTTTTTACACCGACACCGCGATAGCAAAAGAGGTTTCTCAGCA 780
```

QY	761	AACTACAGTGTCTTGCAGAGCAGTGTCTCAGAAGATTTCAAATGATATGGAAGCTCTGGGC	840
DB	761	AACTACAGTGTCTTGCAGAGCAGTGTCTCAGAAGATTTCAAATGATATGGAAGCTCTGGGC	840
QY	841	ATGGAATCAGAGAAATTCATTCTGACCAAGATCAGAGCTTCTTCCACATATAGCACCAAC	900
DB	841	ATGGAATCAGAGAAATTCATTCTGACCAAGATCAGAGCTTCTTCCACATATAGCACCAAC	900
QY	901	TGGTCTCAGAGCGCTCCCGCCTGAACTACCTCGAATGGGTGACTCCCGAGAGAT	960
DB	901	TGGTCTCAGAGCGCTCCCGCCTGAACTACCTCGAATGGGTGACTCCCGAGAGAT	960
QY	961	TCCTACCGAGTGGATACAGGTAGACTTTGGGCGCTTCTGCGCTTTGTACGGCTGTGCGG	1020
DB	961	TCCTACCGAGTGGATACAGGTAGACTTTGGGCGCTTCTGCGCTTTGTACGGCTGTGCGG	1020
QY	1021	ACACAGGCGCATTTTCAAAGAAACCAAGAAAGAAATATTATGTCAAGACTTTACAAGATC	1080
DB	1021	ACACAGGCGCATTTTCAAAGAAACCAAGAAAGAAATATTATGTCAAGACTTTACAAGATC	1080
QY	1081	GACGTTAGCTCCAAACGGGGGAAGACTGGATCACCATATAAGAGGAAACAAACCTGTTCTC	1140
DB	1081	GACGTTAGCTCCAAACGGGGGAAGACTGGATCACCATATAAGAGGAAACAAACCTGTTCTC	1140
QY	1141	TTTCAGGGAACACCAACCCACAGATGTTGGTTGAGTATTTCCCAAAACCACTGATA	1200
DB	1141	TTTCAGGGAACACCAACCCACAGATGTTGGTTGAGTATTTCCCAAAACCACTGATA	1200
QY	1201	ACTCGATTTGTCGGAATCAAGCTGCAACTTGGGAACTGGCATATCTATCAGATTTGAA	1260
DB	1201	ACTCGATTTGTCGGAATCAAGCTGCAACTTGGGAACTGGCATATCTATCAGATTTGAA	1260
QY	1261	GTATACGGTTGCAAGATAACAGATATTCCTGTCTGGAATGTTGGGTATGGTGTCTGGA	1320
DB	1261	GTATATGGTTGCAAGATAACAGATATTCCTGTCTGGAATGTTGGGTATGGTGTCTGGA	1320
QY	1321	CTTATTTCTGACTCCAGATCAATCATCCAAACCAAGGGGACAGAACTGGATGCCCTGAA	1380
DB	1321	CTTATTTCTGACTCCAGATCAATCATCCAAACCAAGGGGACAGAACTGGATGCCCTGAA	1380
QY	1381	AACATCCGCTTGGTAAACAGTCTGCTGGCTGGGCACCTTCCACCGCACCTCATTTCTCTAC	1440
DB	1381	AACATCCGCTTGGTAAACAGTCTGCTGGCTGGGCACCTTCCACCGCACCTCATTTCTCTAC	1440
QY	1441	ATCAATCAGTGGCTCCAAATAGACTGGGGGAGGAGAGATGCTGTGAGGGGCAATCATCAT	1500
DB	1441	ATCAATCAGTGGCTCCAAATAGACTGGGGGAGGAGAGATGCTGTGAGGGGCAATCATCAT	1500
QY	1501	CAGGTTGGGAAGCACCGAGAGAAAGTGTTCATGAGGAAGTTCAAGATCGGTACAGC	1560
DB	1501	CAGGTTGGGAAGCACCGAGAGAAAGTGTTCATGAGGAAGTTCAAGATCGGTACAGC	1560
QY	1561	AACAACGGCTCGGACTGGAAGATGATCATGATGACGCAACGCAAGGCGAAGCTCTTTT	1620
DB	1561	AACAACGGCTCGGACTGGAAGATGATCATGATGACGCAACGCAAGGCGAAGCTCTTTT	1620
QY	1621	GAGGGCAACAAACACTATGATACACTGAGCTGGCGACTTTTTCAGCTCTCTCCACGCGA	1680
DB	1621	GAGGGCAACAAACACTATGATACACTGAGCTGGCGACTTTTTCAGCTCTCTCCACGCGA	1680
QY	1681	TTCAATCAGGATCTACCCGAGAGAGCCACTCATGCGGACTGGGGCTCAGAAATGAGACTG	1740
DB	1681	TTCAATCAGGATCTACCCGAGAGAGCCACTCATGCGGACTGGGGCTCAGAAATGAGACTG	1740
QY	1741	CTGGCTGTGGAATGGAAGCCCTTACAGCTGGACCGACCACTCCCAACGGGAACTTGGTG	1800
DB	1741	CTGGCTGTGGAATGGAAGCCCTTACAGCTGGACCGACCACTCCCAACGGGAACTTGGTG	1800
QY	1801	GATGAATGTGATGACGACGAGGCAACTGCCACAGTGGAAACAGGTGATGATTCACGCTC	1860
DB	1801	GATGAATGTGATGACGACGAGGCAACTGCCACAGTGGAAACAGGTGATGATTCACGCTC	1860

QY	1861	ACAGTGGCACCACTGTGCTGGCCACAGAAAGCCACCGTCTATAGACAGCACCATACAA	1920
Db	1861	ACAGTGGCACCACTGTGCTGGCCACAGAAAGCCCGCGGTCTATAGACAGCACCATACAA	1920
QY	1921	TCAGAGTTTCCCAACATATGGTTTAACTGTGAATTTGGCTGGGGCTCTCACAGACCTTC	1980
Db	1921	TCAGAGTTTCCCAACATATGGTTTAACTGTGAATTTGGCTGGGGCTCTCACAGACCTTC	1980
QY	1981	TGCCACTGGGAACATGACAATCACGCTCAAGTGGAGTGTGTTGACCAGCAAGACG	2040
Db	1981	TGCCACTGGGAACATGACAATCACGCTCAAGTGGAGTGTGTTGACCAGCAAGACG	2040
QY	2041	GGACCCATTCCAGNACACA--CAGGAGATGGCACTTCATCTATTTCCCAAGCTGACGAA	2097
Db	2041	GGACCCATTCCAGNATCACAGCAGGAGATGGCACTTCATCTATTTCCCAAGCTGACGAA	2100
QY	2098	AATCAGAGGGCAAGTGGCTCCGCTGGTCAGCCCTGTGGTTTATTTCCACAGACTCTGCC	2157
Db	2101	AATCAGAGGGCAAGTGGCTCCGCTGGTGGCCCTGTGGTTTATTTCCACAGACTCTGCC	2160
QY	2158	CACCTGCATGACCTTCTGGTATCACATGTCTGGGTCCACGTCGGCACACTCAGGGTCAAA	2217
Db	2161	CACCTGCATGACCTTCTGGTATCACATGTCTGGGTCCACACTCGGCACACTCAGGGTCAAA	2220
QY	2218	CTGGCGCTACCAGAAAGCCAGAGGAGTACGATCAGCTGGTCTGGATGGCCATTTGGACACAA	2277
Db	2221	CTGGCGCTACCAGAAAGCCAGAGGAGTACGATCAGCTGGTCTGGATGGCCCTTTGGACACAA	2280
QY	2278	GGTGACCACCTGGGAAGGAAGGCGTGTCTTGGCTCCCAAGTCTCTGAAACTTTATCAGGTG	2337
Db	2281	GGTGACCACCTGGGAAGGAAGGCGTGTCTTGGCTCCCAAGTCTCTGAAACTTTATCAGGTG	2340
QY	2338	ATTTTCGAGGCGCAATTCGGAAAGAAACCTTGGTGGGATTTGCTGGATGACATTAGT	2397
Db	2341	ATTTTCGAGGCGCAATTCGGAAAGAAACCTTGGTGGGATTTGCTGGATGACATTAGT	2400
QY	2398	ATTAATAACCAATTTTCCAAAGAGATTTGTGCAAAACCCAGCAGACTCGATGATGAAAGAAC	2457
Db	2401	ATTAATAACCAATTTTCCAAAGAGATTTGTGCAAAACCCAGCAGACTCGATGATGAAAGAAC	2460
QY	2458	CCAGAAATTAATTTGATGAACAGCGGAGCCAGCAGGATGAGGAGAGAGGT	2517
Db	2461	CCAGAAATTAATTTGATGAACAGCGGAGCCAGCAGGATGAGGAGAGAGAGGT	2520
QY	2518	GACAAGAACATCTCCAGGAAGCCAGCAATGTGTTGAAGACCTTTAGAACCCTCCTCATC	2577
Db	2521	GACAAGAACATCTCCAGGAAGCCAGCAATGTGTTGAAGACCTTTAGAACCCTCCTCATC	2580
QY	2578	ACCATATAGCCATGAGCGCCCTGGGGGTCTCTCTGGGGGTGTCTGTGGGGTGTGCTGCT	2637
Db	2581	ACCATATAGCCATGAGTGCCTTGGGGGTCTCTCTGGGGGTGTCTGTGGGGTGTGCTGCT	2640
QY	2638	TACTGTCCCTGTTGGCATATAGGATGTCAGAAAAGAAACTTGTCTGCCCTGGAGAACTAT	2697
Db	2641	TACTGTCCCTGTTGGCATATAGGATGTCAGAAAAGAAACTTGTCTGCCCTGGAGAACTAT	2700
QY	2698	AACCTTTGAACCTTGTGGATGTTGAAGTTGAAAAGAACAACTGAATACACAGAGTACT	2757
Db	2701	AACCTTTGAACCTTGTGGATGTTGAAGTTGAAAAGAACAACTGAATACACAGAGTACT	2760
QY	2758	TATTCGGAGGCATGA	2772
Db	2761	TATTCGGAGGCATGA	2775
RESULT 4			
LOCUS	AX133545	2568 bp	DNA
DEFINITION	Sequence 7 from Patent WO0131346.		linear
ACCESSION	AX133545		
VERSION	AX133545.1		
KEYWORDS	human.		
SOURCE			


```
Db 1561 AACACGGCTCGGACTGGAAGATGATTATGGATGACAGCAAAACCAAGGCGAAGCTCTTTT 1620
Qy 1621 GAGGCAACAACTATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCACGCGA 1680
Db 1621 GAGGCAACAACTATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCACGCGA 1680
Qy 1681 TTATCAGGATCTACCCCGAGAGAGCCACATCATGCGCGACTGGGGCTCAGATGGAGCTG 1740
Db 1681 TTATCAGGATCTACCCCGAGAGAGCCACATCATGCGCGACTGGGGCTCAGATGGAGCTG 1740
Qy 1741 CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCGACCACTCCACGCGAACTTGGTG 1800
Db 1741 CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCGACCACTCCACGCGAACTTGGTG 1800
Qy 1801 GATGAATGTGATGACGACCGGCAACTGCCACAGTGGAAACAGTGATGACTTCCAGCTC 1860
Db 1801 GATGAATGTGATGACGACCGGCAACTGCCACAGTGGAAACAGTGATGACTTCCAGCTC 1860
Qy 1861 ACAGGTGGACCACTGTGCTGGCCACAGAAAGCCACCGCTCATAGACAGCACCATCAA 1920
Db 1861 ACAGGTGGACCACTGTGCTGGCCACAGAAAGCCCGCGCTCATAGACAGCACCATCAA 1920
Qy 1921 TCAGAGTTTCCACATATGTTTAACTGTGAATTTGGCTGGGCTCTCAACAGACCTTC 1980
Db 1921 TCAGAGTTTCCACATATGTTTAACTGTGAATTTGGCTGGGCTCTCAACAGACCTTC 1980
Qy 1981 TGCCACTGGGAACATGATCATGCTGACGCTCAAGTGGAGTGTGTGACGACGACG 2040
Db 1981 TGCCACTGGGAACATGATCATGCTGACGCTCAAGTGGAGTGTGTGACGACGACG 2040
Qy 2041 GGACCCATTGAGGATCACA--CAGGAGATGGCAACTTCATCTATTTCCCAAGCTGAGGAA 2097
Db 2041 GGACCCATTGAGGATCACAAGCAGGAGATGGCAACTTCATCTATTTCCCAAGCTGAGGAA 2100
Qy 2098 AATCAGAAGGCAAGTGGCTCGCTGCTGAGCCCTGCTGTTTATTTCCAGAACTCTGCC 2157
Db 2101 AATCAGAAGGCAAGTGGCTCGCTGCTGAGCCCTGCTGTTTATTTCCCAAGAACTCTGCC 2160
Qy 2158 CACTGCTATGACCTTCTGGTATCATGCTCTGGGTCCACGCTCGGCACACATCAGGGTCAAA 2217
Db 2161 CACTGCTATGACCTTCTGGTATCATGCTCTGGGTCCACATCGGCACACATCAGGGTCAAA 2220
Qy 2218 CTGGCTACCAAGAGCCAGAGGATGATGATGCTGCTGGATGGCCATTGACACCAA 2277
Db 2221 CTGGCTACCAAGAGCCAGAGGATGATGATGCTGCTGGATGGCCATTGACACCAA 2280
Qy 2278 GGTGACCACTGGAAGGAGGCGTCTTCTGCTCCACAACTCTGAAACCTTTATCAGGTG 2337
Db 2281 GGTGACCACTGGAAGGAGGCGTCTTCTGCTCCACAACTCTGAAACCTTTATCAGGTG 2340
Qy 2338 ATTTTCAGGGCGGAAATCGGAAAGGAAACCTTTGGTGGGATGCTGTGGATGACATTAGT 2397
Db 2341 ATTTTCAGGGCGGAAATCGGAAAGGAAACCTTTGGTGGGATGCTGTGGATGACATTAGT 2400
Qy 2398 ATTAATAACCAATTTTCAAGAAGATTTGCAAAACCCAGACCTGGATGATAAAGAAC 2457
Db 2401 ATTAATAACCAATTTTCAAGAAGATTTGCAAAACCCAGACCTGGATGATAAAGAAC 2460
Qy 2458 CCAGAAATTAATTTGATGAACAGGAGGACCGCCAGGATACCAAGGTGAAGGAGGT 2517
Db 2461 CCAGAAATTAATTTGATGAACAGGAGGACCGCCAGGATACCAAGGTGAAGGAGGT 2520
Qy 2518 GACAAGAACATCTCCAGGAAGCCAGGCAATGTGTGTAAGACCTTAGA 2564
Db 2521 GACAAGAACATCTCCAGGAAGCCAGGCAATGTGTGTAAGACCTTAGA 2567
```

```
RESULT 5
HSM804921          5421 bp  mRNA  linear  PRI 10-JUL-2002
LOCUS              Homo sapiens mRNA; cDNA DKFZp686M1967 (from clone DKFZp686M1967).
DEFINITION
ACCESSION  AL833608
VERSION    AL833608.1 GI:21734255
```

```
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 5421)
AUTHORS     Duesterhoeft,A., Lauber,J., Mewes,H.W., Weill,B., Amid,C. and
            Wiemann,S.
TITLE       Direct Submission
JOURNAL     Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKFZp686M1967) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cbna/.
FEATURES
            source
            1..5421
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="DKFZp686M1967"
                /tissue_type="cDNA-collection"
                /DH10B; sites sfIIA + SfiIB"
                /dev_stage="adult"
                /dev_stage="adult"
            polyA_signal
            polyA_site
            5381..5386
BASE COUNT  1566 a 1175 c 1218 g 1462 t
ORIGIN
Query Match      84.3%; Score 2338; DB 9; Length 5421;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 2584; Conservative 0; Mismatches 5; Indels 203; Gaps 2;
QY 1 ATGAGAGGGGGGTGCGGCTCTCTGCGCGCTGCTCGCCCTCGCTCGCCCGCGCGGC 60
Db 255 ATGAGAGGGGGGTGCGGCTCTCTGCGCGCTGCTCGCCCTCGCTCGCCCGCGCGGC 314
QY 61 GCTTTTCGCAACCATGATGTGGGATCTATATAAAATTTGAAAGCCCGGGTACCTTACA 120
Db 315 GCTTTTCGCAACCATGATGTGGGATCTATATAAAATTTGAAAGCCCGGGTACCTTACA 374
QY 121 TCTCCTGGTTATCTCTCATTTCTATCACCAAGTGAATAATGCGAATGGCTGATTCAGGCT 180
Db 375 TCTCCTGGTTATCTCTCATTTCTATCACCAAGTGAATAATGCGAATGGCTGATTCAGGCT 434
QY 181 CCGGACCCATACACAGAGAATTATGATCAACTTCAACCCCTCACTTCGATTTGGAGGACAGA 240
Db 435 CCGGACCCATACACAGAGAATTATGATCAACTTCAACCCCTCACTTCGATTTGGAGGACAGA 494
QY 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAATGAAATGACACATTTTAGG 300
Db 495 GACTGCAA-----
QY 301 GGAAGTCTGTGTGGAAGATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTTT 360
Db 503 -----
QY 361 ATCAAAATTTGCTCTGACTACGAAACACATGTCGAGGATTTTCCATACGTTATGAAT 420
Db 503 -----
QY 421 TTCAAGAGAGGTCCTGAATGTTCCAGAGACTACACACACCTAGTGTAGTGAAGTCC 480
Db 503 -----GTCTGAAATGTTCCAGAGACTACACACACCTAGTGTAGTGAAGTCC 552
QY 481 CCGGATTCCTCAAAAATATCCCAACAGCCTTGAATGCACTTATATGCTTTGGGCCA 540
Db 553 CCGGATTCCTCAAAAATATCCCAACAGCCTTGAATGCACTTATATGCTTTGGGCCA 612
```

QY 541 AAGATGTCAGAGATTATCCTTGGAAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAACTCCT 600
Db 613 AAGATGTCAGAGATTATCCTTGGAAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAACTCCT 672
QY 601 CCAGGGGGGATGTTCTGTGCGCTACGACCGGCTAGAAAATCTGGGATGGATTCCCTGATGTT 660
Db 673 CCAGGGGGGATGTTCTGTGCGCTACGACCGGCTAGAAAATCTGGGATGGATTCCCTGATGTT 732
QY 661 GGCCCTCACATTGGCGGTTACTGTGGACAGAAAAACACAGGTCGAATCCGATCCCTCATCG 720
Db 733 GGCCCTCACATTGGCGGTTACTGTGGACAGAAAAACACAGGTCGAATCCGATCCCTCATCG 792
QY 721 GGCATTCTCTCATGTTGTTTTTACACCGACGCGGATAGCAAAAAGAGTTTCTCAGCA 780
Db 793 GGCATTCTCTCATGTTGTTTTTACACCGACGCGGATAGCAAAAAGAGTTTCTCAGCA 852
QY 781 AACTACAGTGTCTTCAGAGCAGTGTCTCAGAAAGATTTCAAATGTTATGGAAGCTCTGGGC 840
Db 853 AACTACAGTGTCTTCAGAGCAGTGTCTCAGAAAGATTTCAAATGTTATGGAAGCTCTGGGC 912
QY 841 ATGGAATCAGGAGAAATTCATTCTGACCAGATCACAGCTTCTTCCAGTATAGCACCAAC 900
Db 913 ATGGAATCAGGAGAAATTCATTCTGACCAGATCACAGCTTCTTCCAGTATAGCACCAAC 972
QY 901 TGGTCTGACAGCGCTCCCGCTGAACTACCTGAGAATGGGTGGACTCCCGGAGAGGAT 960
Db 973 TGGTCTGACAGCGCTCCCGCTGAACTACCTGAGAATGGGTGGACTCCCGGAGAGGAT 1032
QY 961 TCCTACCGAGAGTGATACAGGTAGACTTGGGCTTCTGCGCTTTGTACGCGCTGTGGG 1020
Db 1033 TCCTACCGAGAGTGATACAGGTAGACTTGGGCTTCTGCGCTTTGTACGCGCTGTGGG 1092
QY 1021 ACACAGGGCGCCATTTCAAAAGAAACCAAGAAATATTTATGTCAAGACTTTACAAGATC 1080
Db 1093 ACACAGGGCGCCATTTCAAAAGAAACCAAGAAATATTTATGTCAAGACTTTACAAGATC 1152
QY 1081 GACGTTAGTCTCAACGGGAGACTGGATACCATATAAGAGAGAAACAAACCTGTTCTC 1140
Db 1153 GACGTTAGTCTCAACGGGAGACTGGATACCATATAAGAGAGAAACAAACCTGTTCTC 1212
QY 1141 TTTCAGGAGAAACCAACCCACAGATGTTGTGGTTGAGTATTTCCCAACCACTGATA 1200
Db 1213 TTTCAGGAGAAACCAACCCACAGATGTTGTGGTTGAGTATTTCCCAACCACTGATA 1272
QY 1201 ACTCGATTGTCGGAATCAAGCTCGCACTTGGGAACTGGCATATCTATGAGATTGAA 1260
Db 1273 ACTCGATTGTCGGAATCAAGCTCGCACTTGGGAACTGGCATATCTATGAGATTGAA 1332
QY 1261 GTATACGGTTGCAAGATAACAGATTATCCTTGCTCTGGAATGTTGGGTATGTTCTGGA 1320
Db 1333 GTATACGGTTGCAAGATAACAGATTATCCTTGCTCTGGAATGTTGGGTATGTTCTGGA 1392
QY 1321 CTTATTTCTGACTCCAGATCACATCATCCACCAAGGAGACAGAACTGGATGCTGNA 1380
Db 1393 CTTATTTCTGACTCCAGATCACATCATCCACCAAGGAGACAGAACTGGATGCTGNA 1452
QY 1381 AACATCCGCTGGTAACAGATCGCTCTGCTGGGACATTCACCCGCACTCATTTCCCTAC 1440
Db 1453 AACATCCGCTGGTAACAGATCGCTCTGCTGGGACATTCACCCGCACTCATTTCCCTAC 1512
QY 1441 ATCAATGAGTGCCTCCAAATAGACCTGGGGAGGAGAGATCGTGAGGGGACATCATATT 1500
Db 1513 ATCAATGAGTGCCTCCAAATAGACCTGGGGAGGAGAGATCGTGAGGGGACATCATATT 1572
QY 1501 CAGGTTGGGAACCGGAGAGAACAGGTGTTTCATGAGGAAGTTTCAAGATCGGTTACAGC 1560
Db 1573 CAGGTTGGGAACCGGAGAGAACAGGTGTTTCATGAGGAAGTTTCAAGATCGGTTACAGC 1632
QY 1561 AACAAACGGCTCGGACTGGGAAGATGATCATGGATGACAGCAAAACGCAAGCGGAAGTCTTTT 1620
Db 1633 AACAAACGGCTCGGACTGGGAAGATGATCATGGATGACAGCAAAACGCAAGCGGAAGTCTTTT 1692
QY 1621 GAGGGCAACAACACTATGATACACCTGAGCTGCGGACTTTTTCAGCTCTCTCCACGCGA 1680

Db 1693 GAGGGCAACAACACTATGATACACTAGCTAGCTGCGGACTTTTCCAGCTCTCTCCACGGA 1752
QY 1681 TTCATCAGGATCTACCCGAGAGAGCCACTCATGCGGAGCTGGGGCTCAGAAATGAGCTG 1740
Db 1753 TTCATCAGGATCTACCCGAGAGAGCCACTCATGCGGAGCTGGGGCTCAGAAATGAGCTG 1812
QY 1741 CTGGGCTGTGAAGTGGAAAGCCCTACAGCTGGACCGACACTCCCAAGGGGAATTTGGTG 1800
Db 1813 CTGGGCTGTGAAGTGGAAAGCCCTACAGCTGGACCGACACTCCCAAGGGGAATTTGGTG 1872
QY 1801 GATGAATGTGATGACGACAGCCCAACTGCCACAGTGGCAAGAGGTGATGATTCAGCTC 1860
Db 1873 GATGAATGTGATGACGACAGCCCAACTGCCACAGTGA----- 1911
QY 1861 ACAGGTGGCACCACCTGTCTGGCCACAGAAAGCCCGGTCATAGACAGCACCATAACA 1920
Db 1912 ACAGGTGGCACCACCTGTCTGGCCACAGAAAGCCCGGTCATAGACAGCACCATAACA 1971
QY 1921 TCAGAGTTTCCAAACATATGGTTTTAACTGTGAATTTGGCTGGGGCTCTCACAAAGACCTTC 1980
Db 1972 TCAGAGTTTCCAAACATATGGTTTTAACTGTGAATTTGGCTGGGGCTCTCACAAAGACCTTC 2031
QY 1981 TGCCACTGGGAACATGACAATCACCTGCAGCTCAAGTGGAGTGTGTTGACCAAGCAAGAG 2040
Db 2032 TGCCACTGGGAACATGACAATCACCTGCAGCTCAAGTGGAGTGTGTTGACCAAGCAAGAG 2091
QY 2041 GGACCCATTTCAGGATCACACAGGAGATGGCAACTTTCATCTATTCCCAAGCTGACCAAAAT 2100
Db 2092 GGACCCATTTCAGGATCACACAGGAGATGGCAACTTTCATCTATTCCCAAGCTGACCAAAAT 2151
QY 2101 CAGAAAGGCAAAAGTGGCTCGCTGGTGAGCCCTGTGGTTTTATTCCCAAGACTCTGCCAC 2160
Db 2152 CAGAAAGGCAAAAGTGGCTCGCCCGTGAGCCCTGTGGTTTTATTCCCAAGACTCTGCCAC 2211
QY 2161 TGCATGACCTTCTGTTATCATATGTTCTGGGTCCCAGCTCGGCACACTCAGGGTCAAACTG 2220
Db 2212 TGCATGACCTTCTGTTATCATATGTTCTGGGTCCCAGCTCGGCACACTCAGGGTCAAACTG 2271
QY 2221 CGCTACCAAGACAGAGAGTACCATCAGCTGGTCTGATGGCCATTGGACACCAAGGT 2280
Db 2272 CGCTACCAAGACAGAGAGTACCATCAGCTGGTCTGATGGCCATTGGACACCAAGGT 2331
QY 2281 GACCACTGGAAGGAGGGCTGCTTTGCTCCACAAAGTCTCTGAAACTTTATCAGGTGATT 2340
Db 2332 GACCACTGGAAGGAGGGCTGCTTTGCTCCACAAAGTCTCTGAAACTTTATCAGGTGATT 2391
QY 2341 TTCGAGGGCGAAATCGGAAAAGGAAACCTTGGTGGGATTGCTGTGGATGACATTAGTATT 2400
Db 2392 TTCGAGGGCGAAATCGGAAAAGGAAACCTTGGTGGGATTGCTGTGGATGACATTAGTATT 2451
QY 2401 AATAACCCACATTTTCAAGAGATTGTGCMAAACAGCAGACTCTGATATAAAGAACCCA 2460
Db 2452 AATAACCCACATTTTCAAGAGATTGTGCMAAACAGCAGACTCTGATATAAAGAACCCA 2511
QY 2461 GAAATTAATAATGTAACACAGGGAGCAGCCAGGATACGAAAGGTCAAGGAGAGGTGAC 2520
Db 2512 GAAATTAATAATGTAACACAGGGAGCAGCCAGGATACGAAAGGTCAAGGAGAGGTGAC 2571
QY 2521 AAGAATCATCTCCAGGAAGCCAGGCAATGTTGTAAGACCTTTAGAACCCTCCTCATCACC 2580
Db 2572 AAGAATCATCTCCAGGAAGCCAGGCAATGTTGTAAGACCTTTAGAACCCTCCTCATCACC 2631
QY 2581 ATCATAGCCATGAGCGCCCTGGGGTCTCTGCGGGGCTGTCTGTGGGTCTGTGCTGAC 2640
Db 2632 ATCATAGCCATGAGCGCCCTGGGGTCTCTCTGCGGGGCTGTCTGTGGGTCTGTGCTGAC 2691
QY 2641 TGTGCTGTGTCATAATGGGATGTCAGAAAGAAACTTTGCTGCCCTGGAGAACTATAAC 2700
Db 2692 TGTGCTGTGTCATAATGGGATGTCAGAAAGAAACTTTGCTGCCCTGGAGAACTATAAC 2751
QY 2701 TTTGAACTTGTGGATGGTGAAGTTGAAAAAGACAAACTGAAATACACAGAGTACTTAT 2760

Db 2752 TTTGAACCTTGCGATGGTGAAGTTGAAAAAGACAAACTGAATACACAGACTACTTAT 2811

QY 2761 TCGAGGCGATGA 2772
|||||

Db 2812 TCGAGGCGATGA 2823

RESULT 6
AX147495 2808 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 49 from Patent WO0136632.
ACCESSION AX147495
VERSION AX147495.1 GI:14346652
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2808)
AUTHORS Levine, Z., David, A., Azar, I., Khosravi, R. and Bernstein, J.
TITLE Variants of alternative splicing
JOURNAL Patent: WO 0136632-A 49 25-MAY-2001;
CompuGen Ltd. (IL)
FEATURES
source
Location/Qualifiers
1..2808
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 771 a 667 c 722 g 646 t 2 others
ORIGIN

Query Match 81.5%; Score 2258; DB 6; Length 2808;
Best Local Similarity 90.8%; Pred. No. 0;
Matches 2518; Conservative 2; Mismatches 2; Indels 250; Gaps 1;

QY 1 ATGGAGAGGGGCTGCGCTCCTCTGCGCGTCTGCTCGCCCTGCTCTCGCCCGGCGCGC 60
|||||

Db 287 ATGGAGAGGGGCTGCGCTCCTCTGCGCGTCTGCTCGCCCTGCTCTCGCCCGGCGCGC 346
|||||

QY 61 GCTTTTCGCAACGATGAATGTGGCGTACTATAAAATTTGAAAGCCCGGGTACCTTACA 120
|||||

Db 347 GCTTTTCGCAACGATAAATGTGGCGTACTATAAAATTTGAAAGCCCGGGTACCTTACA 406
|||||

QY 121 TCTCTGGTTATCCTCATTTATACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 180
|||||

Db 407 TCTCTGGTTATCCTCATTTATACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 466
|||||

QY 181 CCGGACCCATACAGAGAATATGATCAACTTCAACCCCTCACTTCGATTGGAGGACAGA 240
|||||

Db 467 CCGGACCCATACAGAGAATATGATCAACTTCAACCCCTCACTTCGATTGGAGGACAGA 526
|||||

QY 241 GACTGCAAGTATGACTACGTTGGAAGTCTTCGATGGAGAAAATGAAAATGGACATTTAGG 300
|||||

Db 527 GACTGCAAGTATGACTACGTTGGAAGTCTTCGATGGAGAAAATGAAAATGGACATTTAGG 586
|||||

QY 301 GGAAGTCTGTGGAAGATAGCCCTCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTT 360
|||||

Db 587 GGAAGTCTGTGGAAGATAGCCCTCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTT 646
|||||

QY 361 ATCAAAATTTGTTCTGACTACGAAACACATGGTGGCAGGATTTTCCATAGCTTATGAAAT 420
|||||

Db 647 ATCAAAATTTGTTCTGACTACGAAACACATGGTGGCAGGATTTTCCATAGCTTATGAAAT 706
|||||

QY 421 TTCAAGAGAGGCTCTGAATTTTCCAGAACTACACAACACCTAGTGGAGTGAATAAGTCC 480
|||||

Db 707 TTCAAGAGAGGCTCTGAATTTTCCAGAACTACACAACACCTAGTGGAGTGAATAAGTCC 766
|||||

QY 481 CCGGATTCCTTGAAATATATCCCAACAGCCTTGAATGACATTAATTTCTTTTGGGCA 540
|||||

Db 767 CCGGATTCCTTGAAATATATCCCAACAGCCTTGAATGACATTAATTTCTTTTGGGCA 826
|||||

QY 541 AAGATGTACAGATTAATCTCTGAATTTGAAGCTTTGACCTGGAGCCTGACCTCAATCT 600
|||||

Db 827 AAGATGTACAGATTAATCTCTGAATTTGAAGCTTTGACCTGGAGCCTGACCTCAATCT 886
|||||

QY 601 CCAGGGGGGATGTTCTGTCGTACGACCGCTAGAAATCTGGGATGGATTCCCTGATGTT 660
|||||

Db 887 CCAGGGGGGATGTTCTGTCGTACGACCGCTAGAAATCTGGGATGGATTCCCTGATGTT 946
|||||

QY 661 GGCCTTCACATTGGGGGTTACTGTGGACAGAAACACAGGTCGAATCCGATCCTCATCG 720
|||||

Db 947 GGCCTTCACATTGGGGGTTACTGTGGACAGAAACACAGGTCGAATCCGATCCTCATCG 1006
|||||

QY 721 GCATTCTCTCCATGTTTATACACCGACAGCGCATAGCAAAAGAGTTTCTCAGCA 780
|||||

Db 1007 GCATTCTCTCCATGTTTATACACCGACAGCGCATAGCAAAAGAGTTTCTCAGCA 1066
|||||

QY 781 AACTACAGTCTCTTGACAGACAGTGTCTCAGAGATTTTCAAAATGTATGGAAGCTCTGGGC 840
|||||

Db 1067 AACTACAGTCTCTTGACAGACAGTGTCTCAGAGATTTTCAAAATGTATGGAAGCTCTGGGC 1126
|||||

QY 841 ATGGAATCAGGAGAAATTCATTCTGACAGATCACAGCTTCTTCCAGTATAGCACCAC 900
|||||

Db 1127 ATGGAATCAGGAGAAATTCATTCTGACAGATCACAGCTTCTTCCAGTATAGCACCAC 1186
|||||

QY 901 TGGTCTGCAGAGCGCTCCCGCTGAACTACCTTGAGAATGGGTGGACTCCCGAGAGGAT 960
|||||

Db 1187 TGGTCTGCAGAGCGCTCCCGCTGAACTACCTTGAGAATGGGTGGACTCCCGAGAGGAT 1246
|||||

QY 961 TCTTACCGAGAGTGGATACAGGTAGACTTGGGCTTCTGCGCTTTTGTACGGCTGTGCGG 1020
|||||

Db 1247 TCTTACCGAGAGTGGATACAGGTAGACTTGGGCTTCTGCGCTTTTGTACGGCTGTGCGG 1306
|||||

QY 1021 ACACAGGGGCCATTTCAAAAGAAACCAAGAGAAATATATGTCAAGACTTACAAGATC 1080
|||||

Db 1307 ACACAGGGGCCATTTCAAAAGAAACCAAGAGAAATATATGTCAAGACTTACAAGATC 1366
|||||

QY 1081 GACGTTAGTCTCAACGGGGAAGACTGGATCACCATAAAGAGGAACAAACCTGTCTC 1140
|||||

Db 1367 GACGTTAGTCTCAACGGGGAAGACTGGATCACCATAAAGAGGAACAAACCTGTCTC 1426
|||||

QY 1141 TTTTACGGGAAACCAACCCACAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
|||||

Db 1427 TTTTACGGGAAACCAACCCACAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1486
|||||

QY 1201 ACTCGATTTGTCGAATCAAGCCTTGCACCTTGGGAACTTGGCATATCTATGAGATTTGAA 1260
|||||

Db 1487 ACTCGATTTGTCGAATCAAGCCTTGCACCTTGGGAACTTGGCATATCTATGAGATTTGAA 1546
|||||

QY 1261 GTATAGGTTGCAAGATAACAGATTATCCTCTCTGGAATGTTGGGTATGTTGTTGTA 1320
|||||

Db 1547 GTATAGGTTGCAAGATAACAGATTATCCTCTCTGGAATGTTGGGTATGTTGTTGTA 1606
|||||

QY 1321 CTTATTTCTGACTCCCGATCACATCATCCAACAGGAGAGACAACTGGATGCTGAA 1380
|||||

Db 1607 CTTATTTCTGACTCCCGATCACATCATCCAACAGGAGACAACTGGATGCTGAA 1666
|||||

QY 1381 ACATCCGCTCGGTAAACAGTCTGCTGGCTGGGCACTTCCACCGACCTCATTCCTTAC 1440
|||||

Db 1667 ACATCCGCTCGGTAAACAGTCTGCTGGCTGGGCACTTCCACCGACCTCATTCCTTAC 1726
|||||

QY 1441 ATCAATAGTGGCTCCAAATAGACCTGGGGAGAGAGAGATCGTGAGGGGCATCATATT 1500
|||||

Db 1727 ATCAATAGTGGCTCCAAATAGACCTGGGGAGAGAGAGATCGTGAGGGGCATCATATT 1786
|||||

QY 1501 CAGGTTGGGAAGCACCGAGAGAAAGGTTTCTATGAGGAAGTTCAAGATCGGGTACAGC 1560
|||||

Db 1787 CAGGTTGGGAAGCACCGAGAGAAAGGTTTCTATGAGGAAGTTCAAGATCGGGTACAGC 1846
|||||

QY 1561 AACACGGCTCGGACTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
|||||

Db 1847 AACACGGCTCGGACTGGAAGATGATGATGATGATGATGATGATGATGATGATGAT 1898
|||||

QY 1621 GAGGCAACAACAATATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCAGCGA 1680
|||||

Db 1899 ----- 1898
|||||

```
QY 1681 TTTCATCAGGATCTACCCGAGAGAGCCACTCATGCGGACTGGGCTCAGAATGGAGCTG 1740
Db 1899 ----- 1898
QY 1741 CTGGGCTGTGAAGTGGAAAGCCCTACAGCTGGACCGACCTCCCAACGGGAACCTTGGTG 1800
Db 1899 ----- 1898
QY 1801 GATGAATGTGTAGCAGCACCAGGCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
Db 1899 ----- 1898
QY 1921 TCAGAGTTTCCAAACATATGGTTTTAACTGTGNAATTTGGCTGGGGCTCTCACAGACCTTC 1980
Db 1957 TCAGAGTTTCCAAACATATGGTTTTAACTGTGNAATTTGGCTGGGGCTCTCACAGACCTTC 2016
QY 1981 TGCCACTGGGAACATGACAATCACGCTGCAGCTCAAGTGGAGTGTCTTGACCAGCAAGACG 2040
Db 2017 TGCCACTGGGAACATGACAATCACGCTGCAGCTCAAGTGGAGTGTCTTGACCAGCAAGACG 2076
QY 2041 GGACCCATTTCAGGATCACACAGGAGATGGCAACTTCATCTATTCCCAAGCTGACGAAAT 2100
Db 2077 GGACCCATTTCAGGATCACACAGGAGATGGCAACTTCATCTATTCCCAAGCTGACGAAAT 2136
QY 2101 CAGAGGGCAAGTGGCTGGCTGTGAGCCGTGTGGTTTATTTCCCAAGCTCTGCCAC 2160
Db 2137 CAGAGGGCAAGTGGCTGGCTGTGAGCCGTGTGGTTTATTTCCCAAGCTCTGCCAC 2196
QY 2161 TGCATGACCTTCTGTGTATCACATGCTGGGTGCCACGCTGGGCACACTCAGGGTCAAACTG 2220
Db 2197 TGCATGACCTTCTGTGTATCACATGCTGGGTGCCACGCTGGGCACACTCAGGGTCAAACTG 2256
QY 2221 CGCTACCAAGACCCAGAGAGTACGATCAGCTGGTCTGGATGGCCATTGGACACCAAGGT 2280
Db 2257 CGCTACCAAGACCCAGAGAGTACGATCAGCTGGTCTGGATGGCCATTGGACACCAAGGT 2316
QY 2281 GACCACTGGAAGGAGGGGGTGTCTTGTCTCCACAAGTCTCTGAAACTTTATCAGGTGATT 2340
Db 2317 GACCACTGGAAGGAGGGGGTGTCTTGTCTCCACAAGTCTCTGAAACTTTATCAGGTGATT 2376
QY 2341 TTCGAGGGCGAAATCGGAAAGAAACCTTGGTGGATTGCTGTGGATGACATTTAGTATT 2400
Db 2377 TTCGAGGGCGAAATCGGAAAGAAACCTTGGTGGATTGCTGTGGATGACATTTAGTATT 2436
QY 2401 AATAACCACTTTCACAGAAGATTGTGCAAAACAGCAGACCTCGATGATGATGATGATGATGAT 2460
Db 2437 AATAACCACTTTCACAGAAGATTGTGCAAAACAGCAGACCTCGATGATGATGATGATGATGAT 2496
QY 2461 GAAATTAATTTGATGAACAGGGAGCAGCGCAGGATACGAGGTGAAGGAGAGGTGAC 2520
Db 2497 GAAATTAATTTGATGAACAGGGAGCAGCGCAGGATACGAGGTGAAGGAGAGGTGAC 2556
QY 2521 AAGAATCTCTCAGGAAGCAGGCAATGTGTGAAGACCTTGAAGCCATCCTCATCACC 2580
Db 2557 AAGAATCTCTCAGGAAGCAGGCAATGTGTGAAGACCTTGAAGCCATCCTCATCACC 2616
QY 2581 ATCATAGCCATGAGCCCTCGGGTCTCTCGGGGGTGTCTGTGGGGTGTCTGTGGGGTGTCTGTAC 2640
Db 2617 ATCATAGCCATGAGYCCCTCGGGTCTCTCGGGGGTGTCTGTGGGGTGTCTGTGGGGTGTCTGTAC 2676
QY 2641 TGTGCTGTGTGCATTAATGGGATGTGAGAAAGAACTTGTCTGTGCCCTGTGAGAACTATATAC 2700
Db 2677 TGTGCTGTGTGCATTAATGGGATGTGAGAAAGAACTTGTCTGTGCCCTGTGAGAACTATATAC 2736
QY 2701 TTTGAACCTGTGTGATGGTGAAGTTGAAAAGACAACTGAATACACAGAGTACTTAT 2760
Db 2737 TTTGAACCTGTGTGATGGTGAAGTTGAAAAGACAACTGAATACACAGAGTACTTAT 2796
QY 2761 TCGGAGGCATGA 2772
```

```
Db 2797 TCGGAGGCATGA 2808
|||||
RESULT 7
AF016296 3471 bp mRNA linear ROD 11-SEP-1997
LOCUS Rattus norvegicus neuropilin mRNA, complete cds.
DEFINITION AF016296
ACCESSION AF016296
VERSION AF016296.1 GI:2367638
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 3471)
AUTHORS Kolodkin,A.L., Levengood,D.V., Rowe,E.G., Tai,Y.T., Giger,R.J. and
Ginty,D.D.
TITLE Neuropilin is a semaphorin III receptor
JOURNAL Cell 90 (4), 753-762 (1997)
MEDLINE 97433085
PUBMED 9288754
REFERENCE 2 (bases 1 to 3471)
AUTHORS Kolodkin,A.L., Levengood,D.V., Rowe,E.G., Tai,Y.T., Giger,R.J. and
Ginty,D.D.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-1997) Neuroscience, Johns Hopkins University
School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205, USA
FEATURES
source
1. 3471
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
144..2913
/note="type I transmembrane protein"
/codon_start=2
/product="neuropilin"
/protein_id="AAC5337.1"
/db_xref="GI:2367639"
/translation="MERGLPLLCLATLALALAGAFRDKCGGKIKIKENPGYLTSPGY
PHYSYSEKCEWLIQAPYORIMINFNPHDLDRDKYDVEIDVGENGGRLMGK
FCGKIAPSPVSSGFFLKFVSDIETHGAGFSIRYIFKRGPECSNYTAPTGVKIS
PGFPEKPNSECTYIIFAPKMSIILEFFSFDLEQSDNPPGVCFRLEIWDGVP
EVGHICRYCGKTPGRISRSSGTLMSVFYVTSIAKEGFSANTSVLQSSISEDFKCM
EALGMESETHSDOITASSOYGTNMSVRSRLNPENGWTPGEDSYREWIOVDLGLLR
FVTAVGQGISKSTKKYVKTIVRDISNGEDWILKEGNKALIIQGNTPNDVVF
GVFPKLITFRVIRKPASWETGISMRFEVYCKITDYPCSGMLGMVSGLISDSQITAS
NQDRNMPENIRLVRTSGTGNALPPSPHPYINELQVLDLGDKEKIVRGVIQGGKHREN
KVFMRKFKIAYSNNGSDMKIMDDSKRKAKEGNNNYDTPELRAFTPLSTRFRIYVP
ERATHSLGLRMELLGCEVEVPTAGPTTNGNPNVDECDQDQANCSGTGDDFQLTGGT
TVLATEKPTIIDSTIOSEFTYGFNCFEGMGSHKTFCHWEHDSHAOLRMVLTSTGTP
IQDHTGDNFIYSOARDENQKGVARLVSPVVSOSAHCQVTFWTHMSGSHVGTURVKL
LQHKPEBYDQLVMVVGQGDHWRKGRVLLHKLSUKLYQVIFEGEIGKGNLGVIAVDII
SINNHIPQEDCAKPTDLDKNTEIKIDETGTPGYEKGKDNKISRPGNVLKTLDP1
LITIIAMSALGVLLGAVGVVLYCACWHNGMSERNLSALENYNPFELVDGVKLRKDKLN
PQSNYSEA"
BASE COUNT 931 a 849 c 912 g 779 t
ORIGIN
Query Match 79.7%; Score 2210.4; DB 10; Length 3471;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 2428; Conservative 0; Mismatches 341; Indels 3; Gaps 1;
QY 1 ATGGAGAGGGGCTCCCGCTCTCTGCGCGCTGCTGCGCCCTGCTGCTCGCCCGCCGGC 60
|||||
Db 145 ATGGAGAGGGGCTCCCGCTCTGCTGTGCGCCAGCTGCGCCCTTGCCCTGCGCCCGCCGGC 204
|||||
QY 61 GCTTTTCCCAACGATGATGCGCATACTATAAAATTAAGAACCCCGGCTACCTTACA 120
|||||
Db 205 GCTTTCCCGCAGCAATAATGCGGGGACTATAAAATTAAGAACCCCGGCTACCTTACA 264
|||||
QY 121 TCTCTGTTTATCTCTCATTTTATCACCCAAAGTGAATAATGCAATGGCTGATTGAGCT 180
```


Db 265 TCTCCGGGTACCTCATTTTACCATCCAAGTGAATGGTAAATCCAAAGCT 324
QY 181 CGGAGCCCATACAGAGAATATGATCAACTTCAACCCCTCACTTCGATTTGGAGGACAGA 240
Db 325 CGGAGCCCATACAGAGAATATGATCAACTTCAACCCCATATTCGATTTGGAGGACAGA 384
QY 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAAAATGAAATGGACATTTTAG 300
Db 385 GACTGCAAGTATGACTATCTGGAAGTATCGATGGAGAAATGAAGTGGCCCTGTGG 444
QY 301 GGAAGTCTCTGGAAGATAGCCCTCTCTCTGTTGTCTTTCAGGGCCATTTCTTTTT 360
Db 445 GGAAGTCTCTGGAAGATAGCCCTCTCACCTGTGTCTTTCAGGGCCATTTCTCTTC 504
QY 361 ATCAAAATTTGCTCTGACTACGAACACATGTTGAGGATTTTCCATAGCTTATGAAT 420
Db 505 ATCAAAATTTGCTCTGACTATGAGACCCACGGGGCAGGATTTTCCATCGCTATGAAT 564
QY 421 TTCAAGAGAGGCTCTGAATTTTCCAGAACTACACAACACCTAGTGGAGTGATAAGTCC 480
Db 565 TTCAAGAGAGGCTCTGAATTTTCCAGAACTATACAGCACCTACTGGAGTGATAAGTCC 624
QY 481 CCGGATTCCTTGAAATATATCCCAACAGCCCTGAATGACATTAATTTCTTTGGGCA 540
Db 625 CCGGATTCCTTGAAATATATCCCAACAGCTTGGAGTGACACCTACATCATCTTTG 684
QY 541 AAGATGTCAGATATATCTGGAATTTGAAAGCTTTGACCTGAGCCCTGACACAATCCT 600
Db 685 AAGATGTCAGATATATCTGAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAG 744
QY 601 CCAGGGGGATGTTCTGCTGCTACGACCGCTAGAAATCTGGGATGGATTCCTGATGT 660
Db 745 CCGGAGGAGTGTCTGCTGCTATGACCGCTGGAGATCTGGGATGGATTTCCCTGA 804
QY 661 GGCCTCACAATTTGGGGTTTACTGTGGACAGAAACACAGGTCGAATCCGATCCTCAT 720
Db 805 GGCCTCACAATTTGGGGTTTACTGTGGGACAGAAATCTCTGGCGGATCGCTCTTCA 864
QY 721 GGCATCTCTCATGTTGTTTACACCGACAGCGGATAGCAAAAGAGTTTCTCAGCA 780
Db 865 GGCATCTCTCATGTTGTTTACACCGACAGCGGATAGCAAAAGAGTTTCTCAGCC 924
QY 781 AACTACAGTGTCTTGAGAGCAGTGTCTCAGAAAGATTTCAAATGATGGAAGCTCTGG 840
Db 925 AACTACAGGCTGTGCAGAGCAGCATCTCTGAAGATTTTCAAGTGTATGGAGCTCTGG 984
QY 841 ATGGAATCAGGAGAAATTCATTTCTGACCAAGATCAGAGCTTCTCCAGTATAGCACC 900
Db 985 ATGGAATCTGGAGATCCATTTCTGACCAAGATCACTGCAATCTTCCAGATGATG 1044
QY 901 TGGTCTGACAGCGCTCCCGCTGAACTACCTGAGAAATGGGTGGACTCCCGGAGAGGAT 960
Db 1045 TGGTCTGTGAGCGCTCCCGCTGAACTACCTGAAACGGGTGGACACAGGAGAGGAC 1104
QY 961 TCCCTACCGAGTGGATACAGTAGACTTTGGCCCTTCTGCGTTTGTTCACGGCTCTGG 1020
Db 1105 TCCCTACCGAGTGGATCCAGGTGGACTTTGGGCCCTCTGCGATTCGTTACTCTGTGG 1164
QY 1021 ACACAGGGGGCCATTTCAAAGAAACCAACCAAGAGAAATATATGTCAGACTTACAAG 1080
Db 1165 ACACAGGGGTGCCATTTCCAAGAAACCAAGAGAAATATATGTCAGACTTACAAG 1224
QY 1081 GAGCTTAGTCTCAACGGGAGACTGGATCACCATAAAGAGGAAACAAACCTGTCTC 1140
Db 1225 GACATCAGCTCCACGGAGAGACTGGATCACCCTGAAGAGGGGAATTAAGCCATATC 1284
QY 1141 TTTTACGGGAAACCAACCAACCAAGATGTTGGTGTGAGTATTTCCCAACCACTGATA 1200
Db 1285 TTTTACGGGAAACCAACCAACCAAGATGTTGGTGTGAGTATTTCCCAACCACTGATA 1344
QY 1201 ACTCCATTTGCCGAATCAAGCTTCAAGCTTGGGAAATGGCAATATCTATGAGATTGAA 1260

Db 1345 ACTCGATTTGTCGGAATCAAACTGCATCCTGGGAACTGGAAATATCTATGAGATTGAA 1404
QY 1261 GTATACGGTTCAAGATAACAGATTATCTGCTCTGGAAATGTTGGTATGTTGTTCTGGA 1320
Db 1405 GTTATGGCTGCAAGATAACAGATTACCTTCTCTGGAAATGTTGGCATGTTGTTCTGGA 1464
QY 1321 CTATTTTCTGACTCCCATCATCATCAACCAAGGAGACAGAACTGGATGTCCTGAA 1380
Db 1465 CTATTTTCTGACTCCCATCATCATCAACCAAGGAGACAGAACTGGATGTCCTGAA 1524
QY 1381 ACATCCCGCTGGTACCACTGCTGCTGGGACATTCACCCGCACTCATTCCTTAC 1440
Db 1525 ACATCCCGCTGGTACCACTGCTGCTGGGACATTCACCCGCACTCATTCCTTAC 1584
QY 1441 ATCAATGAGTGTCTCAATAGACCTGGGGGAGGAGAACTGTTAGGGGCACTCATATT 1500
Db 1585 ATCAATGAGTGTCTCAATGAGCTGGGAGATGAGAAATAGTAGAGGTGTCTCATATT 1644
QY 1501 CAGGTTGGGAACACCGAGAGAAAGTGTTCATGAGGAAGTTTCAAGATCGGGTACAG 1560
Db 1645 CAGGTTGGGAACACCGAGAGAAAGTGTTCATGAGGAAGTTTCAAGATCGCTTACAGT 1704
QY 1561 ACAACGGCTCGGACTGGGAAGATGATCATGATGACAGCAACGCAAGCGGAAGTCTTTT 1620
Db 1705 AACAATGTTCTGACTGGGAATGATCATGATGACAGCAAGCGCAAGGCTAAGTCTTTT 1764
QY 1621 GAGGCAACAACAATATGATACCTGAGCTCGGAGCTTTTCCAGCTCTCTCCACGCA 1680
Db 1765 GAAGCAACAACAATATGACACACTGAGCTCGGGCCCTTTACACCTCTCTCCACAGA 1824
QY 1681 TTCAATCAGGATCTACCCGAGAGAGCCACTCATGCGGACTGGGCTCAGATGAGAGCTG 1740
Db 1825 TTCAATCAGGATCTACCCGAGAGAGCCACATAGTGGGCTCGGACTGAGGATGGAGCTA 1884
QY 1741 CTGGGCTGTGAAGTGGAAAGCTTACAGCTGACAGCTTCCCAACGGGAACTTTGGTG 1800
Db 1885 CTGGGCTGTGAAGTGGAAAGCTTACAGCTGACAGCTTCCCAACGGGAACTTTGGTG 1944
QY 1801 GATGAATGTGATGACAGCCAGGCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
Db 1945 GACGAGTGTGAGATGACCGGCAACTGCCACAGTGGGACAGGTGATGACTTCCAGCTC 2004
QY 1861 ACAGTGGGACCACTGTGTGCTGCCACAGAAAGCCACGGTCAATAGACAGCACTACAA 1920
Db 2005 ACAGGAGGCACTGTCTGCTGGCCACAGAGAAACCACTTATAGACAGCACTACAA 2064
QY 1921 TCAGAGTTTCCAACTATGTTTAACTGTAATTTGGCTGGGCTCTCAGAGACCTTTC 1980
Db 2065 TCAGAGTTTCCCACTATGCTGTTTAACTGCGAGTTTGGCTGGGCTCTCAGAGACATTC 2124
QY 1981 TGCCACTGGGAACATGACAATCACGTGCAAGTGGAGTGTGTTGACCAAGAGAGG 2040
Db 2125 TGCCACTGGGAACATGACAGCCACCGGCACTCAGTGGAGGCTGCTGACCAAGAGAGG 2184
QY 2041 GGACCCATTCAGGATTCACAGGAGATGGCAATTCATCTATTTCCCAAGCTGACGAAAT 2100
Db 2185 GGCCCATTCAGGACCAACAGGAGATGGCAATTCATCTATTTCCCAAGCTGATGAAAT 2244
QY 2101 CAGAAGGCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db 2245 CAGAAGGCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2304
QY 2161 TGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 2305 TGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2364
QY 2221 CGCTACCAAGAGGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db 2365 CACTACCAAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2424
QY 2281 GACCACTGGAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db 2425 GACCACTGGAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2484

Qy 2341 TTCGAGCGCGAAATCGGAAAAAGGAAACCTTGGTGGGATTGCTGTGGATGACATTAGTATT 2400
Dy 2485 TTTGAAGGTGAATTCGGAAAAAGGAAACCTCGTGGGATTGCTGTGGATGATCAGTATT 2544
Qy 2401 AATRACCACATTTACAGAAGATTGTCAAACACGACAGACCTGGATATAAAGAACCCA 2460
Dy 2545 AACAAACCATTTCTCAGGAGGACTGTGCAAAACCAACAGACCTAGATATAAAGAACACA 2604
Qy 2461 GAAATTAATTTGATGAACAGGAGCAGCGCAGGATACGAAGGTGAAGAGAAAGGTGAC 2520
Dy 2605 GAATTTAAATATAGATGAACAGGAGCAGCAGGATATGAA---GAAGGAAAGCGCAGC 2661
Qy 2521 AAGAACATCTCAGGAAGCCAGGCAATGTGTGAAGACCTTTAGAACCCATCCTCATCACC 2580
Dy 2662 AAGAACATCTCAGGAAGCCAGGCAATGTGTGAAGACCTTTAGAACCCATCCTCATCACC 2721
Qy 2581 ATCATAGCCATGAGCGCCTGGGGTCTCTCTGGGGGCTGTCTGGGGTGTGCTGTGATAC 2640
Dy 2722 ATCATAGCCATGAGTGGCCTGGGGTGTCTCTGGGGTGTGCTGTGAGTGTGCTGTGATAC 2781
Qy 2641 TGTGCTGTGGCATAATGGGATGTGCAAAAGAAACTTGTCTGCCCTGGAGAACTATTAAC 2700
Dy 2782 TGTGCTGTGGCATAATGGGATGTGCGAAAGAACTTGTCTGCCCTGGAGAACTATTAAC 2841
Qy 2701 TTTGAACCTGTGGATGGTGTGAAGTTGAAAAAGACAAACTGAATPACACAGAGTACTTAT 2760
Dy 2842 TTTGAACCTGTGGATGGTGTGAAGTTGAAAAAGACAAACTGAATPACACAGAGTACTTAT 2901
Qy 2761 TCGGAGGCATGA 2772
Dy 2902 TCAGAGCGGTGA 2913

RESULT 8
AX305466
LOCUS AX305466 3652 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 217 from Patent WO0188188.
ACCESSION AX305466
VERSION AX305466.1 GI:17644982
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 217 22-NOV-2001;
SCHOOL Juridical Person Nihon University (JP)
FEATURES
Location/Qualifiers
1..3652
source /organism="Mus musculus"
BASE COUNT 965 a 874 c 941 g 872 t
ORIGIN
Query Match 79.5%; Score 2202.4; DB 6; Length 3652;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 2416; Conservative 0; Mismatches 356; Indels 0; Gaps 0;
Qy 1 ATGGAGAGGGGCTGCCCTCTCTGCGCCGTGTGCGCCCTCGTCTCGCCCGCGCGCGC 60
Dy 348 ATGGAGAGGGGCTGCCCTCTCTGCGCCAGCTCGCCCTGCGCCCTCGCCCTCGCGCGC 407
Qy 61 GCTTTTCCACAGATGAATGGGGATACTATAAAATTGAAAGCCCGGGTACTTTACA 120
Dy 408 GCTTTCCCGACGCAAAATGTGGGGACCATATAAAATCGAAACCCAGGGTACTTCACA 467
Qy 121 TCTCTGTGTTATCTTCATCTTATCACCCCAAGTGAAATGCGAATGCGATGCTGATTCAGGCT 180
Dy 468 TCTCCCGGTTACCTTCATCTTACCATCCAAGTGAGAAAGTGTGAATGCTGATGCTCAAGCT 527

Qy 181 CCGGACCCATACCAGAGAAATTTATGATCAACTTCAACCCCTCACTTCGATTTGGAGGACAGA 240
Dy 528 CCGGAAACCCATACCAGAGAAATCAATCAACTTCAACCCACATTTCCGATTTGGAGGACAGA 587
Qy 241 GACTGCAAGTATGACTACGTGGAAAGTCTTCGATGGAGAAATGAAATGGACATTTTAGG 300
Dy 588 GACTGCAAGTATGACTACGTGGAAAGTAAATGATGGGAGAAATGAAGGCGCGCGCTGTGG 647
Qy 301 GGAAGTTCTCTGGAAGATAGCCCTCCTCTGTGTGTCTTCAGGGCCATTTCTTTT 360
Dy 648 GGAAGTTCTCTGGAAGATAGCCCTCCTCTGTGTGTCTTCAGGGCCCTTTCTCTTC 707
Qy 361 ATCAAAATTTGTCTCTGACTACGAACACATGGTGCAGAGATTTCCATACGTTATGAAAT 420
Dy 708 ATCAAAATTTGTCTCTGACTATGAGACACATGGGCGAGGTTTCCATCCGCTATGAATC 767
Qy 421 TTCAAGAGAGGTCTCTGAATGTTCCAGAACTTACAAACACCTAGTGGATGATAAAGTCC 480
Dy 768 TTCAAGAGAGGGCCGAATGTTCTCAGAACTATACAGACCTACTGGAGTGATAAAGTCC 827
Qy 481 CCCGATTTCCCTGAAAAATATCCCAACAGCCTTGAATGACCTTATATTTGTCTTCGCGCA 540
Dy 828 CCTGGTTCCCTGAAAAATACCCCAACCTGCTTGGAGTGCACCTACATCATCTTTGCACCA 887
Qy 541 AAGATGTCAGAGATTAATCTCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAACT 600
Dy 888 AAGATGTCAGAGATTAATCTCTGGAATTTGAAAGCTTTGACCTGGAGCAAGACTCGAATCT 947
Qy 601 CCAGGGGGATGTTCTGCTGCTACGACCGGCTAGAAATCTGGGATGGATTCCTGATGTT 660
Dy 948 CCAGGAGGAATGTTCTGCTGCTATGACCGGCTGGAGATCTGGGATGGATTCCTGGAAGTT 1007
Qy 661 GGCCCTCACATTTGGGCGTTTACTGTGGACAGAAACACACAGGTCGAAATCCGATCTCATCG 720
Dy 1008 GGCCCTCACATTTGGGCGTTTACTGTGGGACAGAAACTCTCTGGCGGATCCGCTCTTCA 1067
Qy 721 GGCATTTCTCCATCGTGTGTTTACACCGACAGCGCGATAGCAAAAGAGTGTTCACAGA 780
Dy 1068 GCGTTCCTATCCATGGTCTTTTACACTGACAGCGCAATAGCAAAAGAGTGTTCACAGCC 1127
Qy 781 AACTACAGTGTCTGACAGAGCAGTGTCTCAGAAGATTTCAAAATGTATGGAAGCTCTGGC 840
Dy 1128 AACTACAGTGTGCTACAGAGCAGCATCTCTGAAGATTTTAAAGTGTATGGAGGCTCTGGC 1187
Qy 841 ATGGAATCAGGAGAAATTCATTTGACACAGATCACAGCTTCTCCACGATATAGCACCAAC 900
Dy 1188 ATGGAATCAGGAGAGATCCATCTCTGATCAGATCAGTCTTACAGATATGGTACCAAC 1247
Qy 901 TGGTCTGACAGCGCTCCCGCTGAACTACCTGAGAAATGGTGGACTCCCGGAGAGAT 960
Dy 1248 TGGTCTGTAGAGCGCTCCCGCTGAACTACCTGAAATGGTGGACTCCAGGAGAGAC 1307
Qy 961 TCCTACCGAGAGTGATACAGGTAGACTTGGGCTTCTGCGCTTTGTACGCGCTGTCGGG 1020
Dy 1308 TCCTACAGAGAGTGSATCCAGGTGGACTTGGGCTTCTGCGATTCGTTACTGCTGTAGGG 1367
Qy 1021 ACAGGGCGCATTTTCAAAAGAACCAAGAAATATTTATGCAAGCTTACAAAGATC 1080
Dy 1368 ACAGGGTGCATTTTCCAGGAACCAAGAAATATTTATGTCAGACTTACAGAGTA 1427
Qy 1081 GACGTTAGCTCCACGGGGAAGACTGGATCACCATTAAAGAGGAAACAAACCTGTTCCTC 1140
Dy 1428 GACATCAGCTCCACGGGAGAGACTGGATCTCCCTGAAAGAGGAAATATAAGCCATTATC 1487
Qy 1141 TTTCAGGGAACACACACCCACAGATGTTGTGGTTCGAGTATTTCCCCCAACCACTGATA 1200
Dy 1488 TTTCAGGGAACACACACCCACAGATGTTGTCTTTAGAGTGTTCCTCCAAACCACTGATA 1547
Qy 1201 ACTCGATTTGTCGAAATCAAGCCTGCAACTTGGGAACTGGCAATCTATCATGAGATTTCAA 1260
Dy 1548 ACTCGATTTGTCGAAATCAAGCCTGATCTCCTGGGAACTGGTATATCTATGAGATTTGAA 1607
Qy 1261 GTATACGGTTGCAAGATAACAGATTTATCCTTGTCTGGAATGTTGGGTATGGTGTCTGGA 1320

Db 1608 GTTATGGCTGCAAGATAACAGATTATCTGCTCTGGAATGTTGGGCATGTTGCTGGA 1667
QY 1321 CTTATTTCTGACTCCAGATACATCATCAACCAAGGAGAGACGAACACTGGATGCTCGAA 1380
Db 1668 CTTATTTAGACTCCAGATACAGATCCATCAACGAGCAGAGAAATGGATGCCAGAA 1727
QY 1381 AACATCCGCTGGTACACAGTGGCTCTGGCTGGGCACTTCCACCCGACCTCATTTCTAC 1440
Db 1728 AACATCCGCTGGTACACAGTGGCTCTGGCTGGGCACTTCCACCCGACCTCATTTCTAC 1787
QY 1441 ATCAATGAGTGGCTCCAAATAGACTGGGGAGGAGAGATCGTAGGGGCATCATCAT 1500
Db 1788 ACCAATGAATGGCTCCAAATGAGCTGGGAGATGAGAAGATAGTAGAGAGGTGTCTCAT 1847
QY 1501 CAGGTGGGAAGCACCGACAGAACAGGTGTTTCATGAGGAAGTTCGAAGATCGGGTACAG 1560
Db 1848 CAGGTGGGAAGCACCGAGAACAGGTGTTTCATGAGGAAGTTCGAAGATCGGGTACAG 1907
QY 1561 AACACGGCTCGGACTGGAAGATGATCATGGATGACAGCAACGCAAGCGAAGTCTTTT 1620
Db 1908 AACATGGCTCTGACTGGAAACTATCATGGATGACAGCAAGCGCAAGGCTAAGTCGTT 1967
QY 1621 GAGGCAACAACTATGATACACCTGAGCTGGGACTTTTCCAGCTCTCTCCAGCGA 1680
Db 1968 GAAGCAACAACTATGATACACCTGAGCTGGGACTTTTCCAGCTCTCTCCAGAGG 2027
QY 1681 TTTCATCAGGATCTACCCCGAGAGACCTCATGCGGACTGGGCTCAGAAATGGAGCTG 1740
Db 2028 TTTCATCAGGATCTACCCCGAGAGACCTCATGCGGACTGGGCTCAGAAATGGAGCTG 2087
QY 1741 CTGGGCTGGAAGTGAAGCCCTACAGCTGGACGACCACTCCCAACGGGAACTTGGTG 1800
Db 2088 CTGGGCTGGAAGTGAAGCCCTACAGCTGGACGACCACTCCCAACGGGAACTTGGTG 2147
QY 1801 GATGAATGATGACGACGACCCCACTCCACAGTGGAAAGGTGATGACTTCCAGCTC 1860
Db 2148 CATGAGTGTGACGAGACGACGACCACTCCACAGTGGGACAGGTGATGACTTCCAGCT 2207
QY 1861 ACAGTGGCACCACCTGTGCTGCCACAGAAAGCCAGCTCATAGACAGCACCATAAA 1920
Db 2208 ACAGGAGGACCACTGTCTGCCACAGAGAACCACTATATAGACAGCACCATAAA 2267
QY 1921 TCAGAGTTTCAACATATAGTTTAACTGTGAATTTGGCTGGGCTCTCCAAAGACCTTC 1980
Db 2268 TCAGAGTTTCCGACATACGTTTTAACTGCGAGTTTGGCTGGGCTCTCCAAAGACAT 2327
QY 1981 TCCCACTGGGACATCACAATCAGCTGCAGCTCAAGTGGAGTGTGTGACGACAGAGC 2040
Db 2328 TCCCACTGGGACATGACAGCCATGCACAGCTCAGGTGGAGTGTGTGACGACAGACA 2387
QY 2041 GGACCCATTTCAGGATCACACAGAGATGGCAACTTCATCTATTCACCAAGCTCAGCAAA 2100
Db 2388 GGCCGATTTCAGGATCACACAGAGATGGCAACTTCATCTATTCACCAAGCTCAGCAAA 2447
QY 2101 CAGAGGGCAAGTGGCTCGCTGGTGGAGCCCTGTGGTTTATTCACAGAACTCTGCCAC 2160
Db 2448 CAGAAAGGCAAGTGGCTGGTGGAGCCCTGTGGTCTATTCACAGAGCTCTGCCAC 2507
QY 2161 TCCATGACCTTCTGTATACATGTTCTGGTCCGAGCTCGGACACACTCAGGGTCAACTG 2220
Db 2508 TGTATGACCTTCTGTATACATGTTCTGGTCCGAGCTCTCATGTGGGTACACTGAGGGTCA 2567
QY 2221 CGCTACCAAGACGAGAGTACGATCAGCTGGTCTGTGATGGCCATTTGGACACCAAGGT 2280
Db 2568 CGCTACCAAGACGAGAGTATGATCACTGGTCTGGATGGTGGTGGGACCAAGGA 2627
QY 2281 GACCACTGGAAGAGGGCGTCTGTGCTCCAAAGTCTCTGAAACTTTATCAGGTGAT 2340
Db 2628 GACCACTGGAAGAGGGCGTCTGTGCTGCAAAATCTCTGAAACTTATCAGGTGAT 2687
QY 2341 TTCAGGGCGAAATCGGAAAGAAACCTTGGTGGGATGCTGTGGATGACGATGATGAT 2400

Db 2688 TTTGAAGTGAAATCGGAAAGGAAACCTTGTGTGGAATTTGCTGTGGATGATATCATGAT 2747
QY 2401 AATAACACATTTCAAGAAAGATTGTGCAAAACACGACAGCTGTGATAAAAAGAACCCA 2460
Db 2748 AACACCATATTTCTCAGGAAGACTGTGCAAAACACGACAGCTAGTAAAGAAACACA 2807
QY 2461 GAAATTTAAATTTGATGAACACAGGAGCAGCCAGGATACGAAGGTGAAGGAGAGGTGAC 2520
Db 2808 GAAATTTAAATTTGATGAACACAGGAGCAGCTCCAGGATATGAAGGAGAGGAGGTGAC 2867
QY 2521 AGACATCTCCAGGAGCCAGGCAATGTTTGAAGACCTTGAAGACCACTCTCATCACC 2580
Db 2868 AGACATCTCCAGGAGCCAGGCAATGTTTGAAGACCTTGAAGACCACTCTCATCACC 2927
QY 2581 ATCATAGCCATCAGGCGCTGGGCTCCCTGGGGCTGTCTGGGGTCTGTGGTGTGAC 2640
Db 2928 ATCATAGCCATCAGGCGCTGGGCTGGAGTACHCTTGGGTGAGTCTGTGGAGTGTGCTGTAC 2987
QY 2641 TCTGCTGTGGCATAATGGGATGTCAGAAAGAAACTTGTCTGCCCTGGAGAACTATAAC 2700
Db 2988 TCTGCTGTGGCATAATGGGATGTCAGAAAGAACTATCTGCCCTGGAGAACTATAAC 3047
QY 2701 TTTGAACCTGTGGATGGTGAAGTTGAAAAAGACAACTGAATACACAGAGTACTTAT 2760
Db 3048 TTTGAACCTGTGGATGGTGAAGTTGAAAAAGATAAACTGAACCCACAGAGTAATTAC 3107
QY 2761 TCGGAGGCGATGA 2772
Db 3108 TCAGAGGCGTGA 3119

RESULT 9
MUSMONE
LOCUS MUSMONE 3652 bp mRNA linear ROD 01-FEB-2000
DEFINITION Mouse mRNA for neuropeilin, complete cds.
ACCESSION D50086
VERSION D50086.1 GI:1783300
KEYWORDS neuropeilin; cell surface protein.
SOURCE Mus musculus (strain:Balb/c) Embryo Brain cDNA to mRNA, clone:ML71.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (sites)
AUTHORS Kawakami,A., Kitsukawa,T., Takagi,S. and Fujisawa,H.
TITLE Developmentally regulated expression of a cell surface protein,
neuropeilin, in the mouse nervous system
J. Neurobiol. 29 (1), 1-17 (1996)
JOURNAL 96353149
MEDLINE 2 (bases 1 to 3652)
REFERENCE AUTHORS Kawakami,A.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3652)
AUTHORS Kawakami,A.
TITLE Direct Submission
JOURNAL Submitted (07-APR-1995) Atsushi Kawakami, Nagoya University,
Biological Science; Furo-cho, Chikusa-ku, Nagoya, Aichi 464-01,
Japan (E-mail:i45526@nuc.cc.nagoya-u.ac.jp, Tel:81-52-789-2981,
Fax:81-52-789-2979)
COMMENT On Jan 17, 1997 this sequence version replaced gi:1199488.
SEQUENCE updated (15-Jan-1997) by: Atsushi Kawakami.
FEATURES
source
1. 3652
/organism="Mus musculus"
/strain="Balb/c"
/db_xref="taxon:10090"
/clone="ML71"
/tissue_type="Brain"
/dev_stage="Embryo"
348. .3119
/codon_start=1
/product="neuropeilin"
/protein_id="BAA08789.1"
/db_xref="GI:1783301"

CDS

/translation="MERGLPLLCATLALALAGAFRSDKGGTIIKINPGYLPSPGY
PHSYHSEKEMWLPQAPYQRIIINPHFDLRDCYDYVEVIGENEGGRLWGK
FCGKIAPDSYSGPFLFIKIVSDYETHAGFSRYEIKRGPESQNYTAPTGVIKS
PGPEKYNCLCEYIIIFAPKMSIILEFSEFDLEQSDNPPGGMFCRYDRLEIWDGPP
EVGPHIGYCGOKTPGIRSRSSGVLSMVFYDLSAIAKEGFSANYSVLQSSISEDFKCM
EALMESGEIHSDDITASSQYGTWNVSERSLANTPENGWTPGEDSYKEWQVLDGLLR
FVAVGTQGLAISRETKKKYVYKTVRVDISSNGEDWIILKEGNKALIFQGNTPDVL
GYKSPILITFRVIRKPYSWETGISMRFEVYGGCKITDYPCSMGLGMVSLIDSQITAS
NQADRNMPENIRLVTSGTGWALPSPHPYTNELQVLDGLDEKIVRGVIIOGGKHREN
KVFMRKIKIAYSNGSDWKIMDDSKRKAKEGFSNNDYDPELRTFSPLETRIYP
ERATHSGLGLRMLLCEVEAPTAGPTPNGNPVHECEDDOANCHSGTGDQFLTGCT
TVLATEPTIIDISIQSEBPYIYCNFEGWGSHTFCHWEHDSHAQLRWSVLTSGTP
IQDHTGDNFIYSOADENQKVARLSPVYVYSSAHCMFTFWHMSGHVSGLTVLRKL
RYKRPEDYDLVMVVGQDHMKREGLVLLKSLKLYQVIFEGEIGKNGLIQVADDI
SINHHISEQCAKPTDLDKKNTEIKIDETGSTPGYEGEGEDKNISIRKPGNVLTLP
ILITIIAMSLGLVLGAVGVLYLCACWHNGMSRNLSALENYFELVDGVKLLKDKL
NPOSNYSEA"
348..410
/note="mouse neuroligin signal peptide"
411..3116
/product="mouse neuroligin"
BASE COUNT 965 a 874 c 941 g 872 t
ORIGIN

Query Match 79.5%; Score 2202.4; DB 10; Length 3652;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 2416; Conservative 0; Mismatches 356; Indels 0; Gaps 0;
QY 1 ATGGAGAGGGGCTCCGCTCTCTGGCGGTGCTGCGGCTGCTGCTGCGCGCGCGCGG 60
DB 348 ATGGAGAGGGGCTCCGCTCTCTGGCGGTGCTGCGGCTGCTGCTGCGCGCGCGG 407
QY 61 GCTTTTCCGACGATGAATGGCCATATATAAAATTTGAAAGCCCGGGTACCTTACA 120
DB 408 GCTTTCCGACGACAAATGGCGGGACCATATAAAATTCGAAACCCAGGGTACCTACA 467
QY 121 TCTCTGGTTATCTCTATATACCCAAAGTGAATAATGCGAATGGCTGATTCAGGCT 180
DB 468 TCTCCGGTTACCTCTCACTTACCATCCAAGTGAGAAGTGTGAATGGCTAATCCAAGCT 527
QY 181 CCGGACCCATACGAGAAATATGATCAACTCAACCTCACTTCGATTTGGAGACAGA 240
DB 528 CCGGAACCTTACCAGAAATCATATACTTCAACCCACATTTCCGATTTGGAGACAGA 587
QY 241 GACTGCAAGTATGACTAGCTGGAAGTCTCGATGGAGAAATGAAATGGACATTTAGG 300
DB 588 GACTGCAAGTATGACTAGCTGGAAGTATGATGGGAGAAATGAAGCGCGCGCTGTGG 647
QY 301 GGAAGTTCTGTGGAAAGATAGCCCGCTCTCTGTGTGTCTTCAGGGCCATTTCTTTT 360
DB 648 GGAAGTTCTGTGGAAAGATGACACTTCTCTGTGTGTCTTCAGGGCCCTTTCTCTTC 707
QY 361 ATCAATTTGTCTGACTACGAAACACATGGTGAGAGATTTTCCATAGTTATGAAAT 420
DB 708 ATCAATTTGTCTGACTATGAGACATGGGGCAGGGTTTCCATCCGCTATGAAATC 767
QY 421 TTCAAGAGAGGTCCTGAAATGTTTCCAGAACTACACACACCTAGTGGAGTGATAAGTCC 480
DB 768 TTCAAGAGAGGCGCGAATGTTCTCAGAACTATACAGACACTACTGGAGTGATAAGTCC 827
QY 481 CCGGATTTCCCTGAAAAATATCCCAACAGCCCTTGAATGACACTATATATTGCTTTTCGCCA 540
DB 828 CCGGATTTCCCTGAAAAATATCCCAACCTGCTTGGAGTGACACCTATCTTTGCACCA 887
QY 541 AGATGTGAGAGATTTATCTGGAATTTTGAAGCTTTGACCTGGAGCCCTGACTCAATCT 600
DB 888 AGATGTGAGAGATTTATCTGGAATTTTGAAGCTTTGACCTGGAGCAAGACTCGAATCT 947
QY 601 CCAGGGGGATGTTCTGTGCTGCTACGACCGGCTAGAAAATCTGGGATGGATTTCCCTCATGTT 660
DB 948 CCGGAGGAATGTTCTGTGCTATGACCGGCTGGAGATCTGGGATGGATTTCCCTGAAGTT 1007
QY 661 GSCCCTCACATTTGGCGGTTACTGTGGACAGAAAACACAGGTGCGAATCCGATCTCATCG 720

DB 1008 GSCCCTCACATTTGGCGGTTATTGTGGGAGAAAATCTCTGCGGATCCGCTCTCTTCA 1067
QY 721 GGCATTCTCTCCATGGTTTTTACACCCACAGCGGATAGCAAAAAGAGGTTTCTTCACGA 780
DB 1068 GCGGTTTATCCATGGTCTTTTACACTGACAGCGCAATAGCAAAAAGAGGTTTCTTCAGCC 1127
QY 781 AACTACAGTGTCTTGACAGAGCAGTGTCTCAGAAATTTCAAAATGTATGAAAGCTCTGGC 840
DB 1128 AACTACAGTGTCTTACAGAGCAGCATCTCTGAAGATTTTAAAGTGTATGAGGCTCTGGC 1187
QY 841 ATGGAATCAGGAGAAATTCATTTCTGACCAGATCAGAGCTTCTTCCCATGTATAGCAACAC 900
DB 1188 ATGGAATCAGGAGATCCATTTCTGATCAGATCAGTGTCTTCCAGATGTGGTACCAAC 1247
QY 901 TGGTCTGACAGCGCTCCCGCTGAACTACCTGAGAAATGGTGGGACCTCCCGGAGAGAT 960
DB 1248 TGGTCTGTAGAGCGCTCCCGCTGAACTACCTGAAATGGTGGACTCCAGGAGAGAC 1307
QY 961 TCCTACCGAGAGTGTATACAGTGTAGCTTTGGGCTTCTGCGCTTTGTCCAGGCTGTCGG 1020
DB 1308 TCCTACAAGAGTGTGATCCAGGTGGACTTGGGCTCTCTGCGATTCGTTACTGCTGAGG 1367
QY 1021 ACACAGGCGCATTTCAAAAGAAACCAAGAAATATTTATGTCAAGACTTTACAAGATC 1080
DB 1368 ACACAGGCTGCGATTTCCAAAGGAAACCAAGAAATATTTATGTCAAGACTTTACAGATA 1427
QY 1081 GACGTTACTCTCAAGGGGAGAGCTGATCACCATAAAGAAAGGAAACAAACCTGTTCTC 1140
DB 1428 GACATAGCTCCACGGAGAGGACTGGATCTCCCTGAAAGAGGAAATAAAAGCCATATC 1487
QY 1141 TTTCAAGGAAACACCAACCCACAGATGTTGTGTGTCAGATTTCCCAAAACCACTGATA 1200
DB 1488 TTTCAAGGAAACACCAACCCACAGATGTTGTCTTAGAGTTTCTCCAAACCACTGATA 1547
QY 1201 ACTCGATTTGTCGGAATCAAGCTTCAACTTTGGGAACTGGCATATCTATGAGATTTGAA 1260
DB 1548 ACTCGATTTGTCGGAATCAAACTGTATCTCTGGGAACTGGTATATCTATGAGATTTGAA 1607
QY 1261 GTATACGTTGCAAGATTAACAGATTTATCTTCTCTGCTGGAATGTTGGGTATGGTCTCGA 1320
DB 1608 GTTATGCTGCAAGATTAACAGATTTATCTTCTCTGCTGGAATGTTGGGATGGTGTCTCGA 1667
QY 1321 CTATTTCTGACTCTCCAGATCACATCATCAACCAAGAGAGACAGAACTGTGATGCTGAA 1380
DB 1668 CTATTTGACTCTCCAGATTTACAGCATCCAATCAAGCCGACAGAAATTTGATGCCAGAA 1727
QY 1381 AACATCCGCTGGTAACCAAGTCTGCTGGTGGGCACTTTCCACCCGCACTCATTTCTCTAC 1440
DB 1728 AACATCCGCTGGTGACCACTGCTACCGCTGGGCACTGCCACCTCACCCACCCACATAC 1787
QY 1441 ATCAATGAGTGGCTCCAATAGACCTGGGGAGGAGAGATCGTGCAGGGCATCATCAT 1500
DB 1788 ACCAATGAATGGCTCCAAGTGGACCTGGGAGATGAGAAGATAGTAAGAGGTGTCATCAT 1847
QY 1501 CAGGTGGGAAGCAGCAGAGAGAAAGTGTTCATGAGAGAGTTCAAGATCGGGTACAGC 1560
DB 1848 CAGGTGGGAAGCAGCAGAGAGAAAGTGTTCATGAGAGAGTTCAAGATCGCCTATAGT 1907
QY 1561 AACACGCTCGGATGGAAGATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1908 AACATGCTCTGACTGGAAACTATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1967
QY 1621 GAGGCAACAACTATGATACACTGACCTGAGCTGGCGGACTTTTCCAGCTCTCTCCACGGA 1680
DB 1968 GAAGCAACAACTATGACACACTGAGCTTGGAGCTTTTCCAGCTCTCTCTCCACAGG 2027
QY 1681 TTCAATCAGGATCTACCCCGAGAGAGCCACTCATGCGGACTGGGGCTCAGAAATGAGAGCTG 1740
DB 2028 TTCAATCAGGATCTACCCCTGAGAGAGCCACACAGCTGGGCTTGGGCTCAGGATGGAGCTA 2087
QY 1741 CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCGACCTCCACAGCGGAACCTTGTG 1800

|||||
Db 238 GACTCAAGTATGACTGTGGAAGTGATCGATGAGAGAATGAAGGTGGCGCTGTGG 297
QY 301 GGAAGTCTGTGGAAAGATAGCCCTCTCTGTGTGTCTTCAGGGCCATTTCTTTT 360
Db 298 GGAAGTCTGTGGAAAGATAGCCCTCTCTGTGTGTCTTCAGGGCCATTTCTTTT 357
QY 361 ATCAAAATTTGTCTGACTACAAAACACATGCTGAGGATTTTCCATACGTTATGAAAT 420
Db 358 ATCAAAATTTGTCTGACTATGAGACCCACGGGAGGATTTTCCATCCGCTATGAATC 417
QY 421 TTCAGAGAGGTCCTGAATTTCCAGAACTACAAACACCTAGTGGAGTGATAAGTCC 480
Db 418 TTCAGAGAGGCGCGGAATTTCTCAGAACTATACAGACCTACTTGGAGTGATAAGTCC 477
QY 481 CCCGATTCCTTGAAAAATATCCCAACAGCCTTGAAATGCACCTATATCTCTTTGGCCA 540
Db 478 CTGGGTTCCCTTGAAAAATACCCCAACAGCTTGGAGTGACCTATCATCTTTGCACCA 537
QY 541 AAGATGTCAGAGATTTATCCTTGGAAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAACTCT 600
Db 538 AAGATGTCAGAGATTAATCCTAGAGTTTGAAGTTTGAACCTGGAGCAAGACTCAAACTCT 597
QY 601 CCAGGGGGATGTTCTGTGCTACAGCCGCTAGAAATCTGGGATGGATTCCTGATGTT 660
Db 598 CCCGAGGAATGTTCTGTGCTATGACCGGCTGGAGATCTGGGATGGATTCCTGAAAGTT 657
QY 661 GGCCCTCACATTTGGCGTTTACTGTGCAGACAAAACACAGGTCGAATCCGATCTCTCATCG 720
Db 658 GGCCCTCACATTTGGCGTTTACTGTGGCGAAGAAATCTCTGGCCGGATCCGCTCTCTTCA 717
QY 721 GCATTTCTTCATCGTTTATTTACACGACAGCGGATAGCAAAAGAGTTTTCACGCA 780
Db 718 GGCATTTATCCATGTTCTTCTACACTGACACGCAATAGCAAGAGTTTTCACGCC 777
QY 781 AACTACAGTCTTCGAGAGAGTGTCTCAGAAGATTTCAAAATGTATGGAAGCTCTGGGC 840
Db 778 AACTACAGGCTCTCGAGAGAGCATCTCTGAAGATTTCAAGTGTATGAGGCTCTGGGC 837
QY 841 ATGGAATCAGGAATAATTCATTTGACCAGATCACAGCTTCTCCAGTATAGCACCAAC 900
Db 838 ATGGAATCTGGAGAGATTCATTTGACCAAGATCACTGCATCTTCCAGTATGTTACCAAC 897
QY 901 TGGTCTGCAGAGCGTCCCGCTGAACCTACCCTGAGATGGGTGGACTCCCGGAGAGAT 960
Db 898 TGGTCTGTGAGCGTCCCGCTGAACCTACCCTGAAACGGGTGGACACAGAGAGAC 957
QY 961 TCCCTACAGAGTGGATACAGTACAGTACGTTGGGCTTCTGCGCTTTGTACGCGCTGTGGG 1020
Db 958 TCCCTACAGGAGTGGATCCAGTGGACTTGGGCTTCTGCGATTCGTTACTGCTGTGGG 1017
QY 1021 ACACAGGGCGCATTTTCAAAAGAACCAAGAGAAATATATGTCAAGACTTTACAAGTA 1080
Db 1018 ACACAGGGTGCATTTTCAAGGAACCAAGAGAAATATATGTCAAGACTTTACAAGTA 1077
QY 1081 GACGTAGTCTCAACGGGGAAGACTGGATCACCATTAAGAGAGAAACAAACCTGTTCTC 1140
Db 1078 GACATCAGCTCAACGGGAGAGACTGGATCACCCCTGAAGGAGGAAATAAGGCCATATC 1137
QY 1141 TTTACGGGAAACACCAACCCACAGATGTTGTGGTTGAGTATTCCTCCCAACCACTGATA 1200
Db 1138 TTTACGGGAAACCAACCAATCCCAAGATGTTGTCTTTGGAGTTTTCCTCCCAACCACTGATA 1197
QY 1201 ACTCGATTTGTCGGAATCAAGCTGCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1260
Db 1198 ACTCGATTTGTCGGAATCAAACTGCATCTCTGGGAACTGGAAATATCTATGAGATTTGAA 1257
QY 1261 GTATACGTTGCAAGATAACAGATTTATCCTGTCTGGAATTTGGGATGTTGTCTGGA 1320
Db 1258 GTTATGCTGCAAGATAACAGATTTACCTTCTCTGGAATTTGGGATGTTGTCTGGA 1317
QY 1321 CTTATTTCTGACTCCAGATCATCATCTCCAAACCAAGGAGACAGAACTGGATGCTGAA 1380
|||||

Db 1318 CTTATTTCAGACTCCAGATTTACAGCATCAACAAAGGAGACAGAACTGGATGCCAGAA 1377
QY 1381 AACATCCGCTCGTAACAGTGGCTGTGGTGGCACTTTCCACCGCACCTCATTTCTCTAC 1440
Db 1378 AACATCCGCTCGTAACAGTGGAAACCGCTGGGCTTCCACCCCTCACCCACCCATAC 1437
QY 1441 ATCAATGATGCTCCAAATAGACCTGGGGAGGAGAAAGATCGTAGGGGCACTCATCAT 1500
Db 1438 ATCAATGATGCTCCAAATGACCTGGGAGATGAGAAGATAGTAAGAGGTGTCATCAT 1497
QY 1501 CAGGTGGGAAGCACCGGAGAGAAAGGTGTTTCATGAGGAAGTTCAAGATCGGGTACAGC 1560
Db 1498 CAAAGTGGGAAGCACCGGAGAAACAAAGTGTTCATGAGGAAGTTCAAGATCCCTACAG 1557
QY 1561 AACAAAGCTCGGACTGGAAGATGATCATGATGACAGCAACGAAAGCGGAAGTCTTTT 1620
Db 1558 AACATGTTCTGACTGGAAATGATCATGATGACAGCAAGCGCAAGGCTAAGTCTTTT 1617
QY 1621 GAGGCAACAAACACTATGATACACTGAGCTGCGGACTTTTTCAGCTCTCTCCACGGGA 1680
Db 1618 GAAGCAACAAACACTATGACACACTGAGCTCCGGGCTTTTACACCTCTCTCCACAAGA 1677
QY 1681 TTCATCAGGATCTCCCGAGAGAGCCACTCATGCGGACCTGGGCTCAGAAATGAGCTG 1740
Db 1678 TTCATCAGGATCTCCCGAGAGAGCCACACATAGTGGGCTCGGACTGAGGATGGAGCTA 1737
QY 1741 CTGGCTGTGAAGTGGAAAGCCCTTACAGCTGGACCGACACACTCCCAAGCGGAACCTTGGTG 1800
Db 1738 CTGGGCTGTGAAGTAGAAGTGCCTACAGCTGGACCCACGACACCAATGGGAACCCCTG 1797
QY 1801 GATGAATGTGATGACGACAGGCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
Db 1798 GACGAGTGTGAGGATGACAGGCCAACTGCCACAGTGGCACAGGTGATGACTTCCAGCTC 1857
QY 1861 ACAGTGGCACCACTGTGCTGGCCACAGAAAGCCCGCTCATAGACAGCACCATACAA 1920
Db 1858 ACAGGAGGCACCACTGTCTTGGCCACAGAAAGCCCACTATATAGACAGCACCATCCAA 1917
QY 1921 TCAGAGTTTCCAAACATATGGTTTAACTGTGAATTTGGCTGGGCTCTCACAAAGACTTTC 1980
Db 1918 TCAGAGTTTCCCGACATACGGTTTAACTGCGAGTTTGGCTGGGCTCTCACAAAGACTTTC 1977
QY 1981 TGCCACTGGGAACATGACAATCACGTGACGTCAAGTGGAGTGTGTTGACACAGCAAGAG 2040
Db 1978 TGCCACTGGGAACATGACAGCCACGCGAGCTCAGTGGAGGTGCTGACACAGCAAGAG 2037
QY 2041 GGACCCATTCAGGATCACACAGGAGATGGCAACTTCATCTATTCCCAAGCTCACGAAAT 2100
Db 2038 GGGCCCATTCAGGACCCACACAGGAGATGGCAACTTCATCTATTCCCAAGCTGATGAAAT 2097
QY 2101 CAGAAGGGCAAGTGGCTCGCCTGGTGGAGCCCTGTGGTTTATTTCCAGAACTCTGCCAC 2160
Db 2098 CAGAAGGCAAGTAGCCGCTGGTGGAGCCCTGTGGTCTATTCCAGAGTTCTGCCAC 2157
QY 2161 TGCATGACCTTCTGATATCATGTCTGGTCCCGCTGCGGACACTCAGGCTCAAACTG 2220
Db 2158 TGCATGACCTTCTGATATCATGTCTCGGCTCTCATGTGGGTACACATGAGGCTCAAACTG 2217
QY 2221 CGCTACCAAGACAGAGGAGTACCATCAGCTGGTCTGGATGGCCATTTGGACACCAAGGT 2280
Db 2218 CACTACCAAGACCCAGAGGAATATGATCAACTGGTCTGGATGGTGGGCGCCCAAGGA 2277
QY 2281 GACCACTGGAAGGAAGGCGTGTCTTGGTCCACAAAGTCTCTGAAACTTTTATCAGGTGAT 2340
Db 2278 GACCACTGGAAGGAAGGCGTGTCTTGGTGCACAAATCTCTGAAACTGTATCAGGTAT 2337
QY 2341 TTCAGGCGGAATCGGAAAGAAACCTTGGTGGGATTTGGTGTGATGACATTTAGTAT 2400
Db 2338 TTTGAAGTGAATCGGAAAGAAACCTCGGTGGGATTTGCTGTGATGATATCAGTAT 2397
QY 2401 AATAACCACATTTACAGAGAGATTTGTCCAAACACAGACCTGGATTAAGAAACCA 2460
Db 2398 AACAAACCAATTCCTCAGGAGGACTGTGCAAACCAACAGACCTAGATTAAGAAACCA 2457
|||||

QY 1261 GTATACGGTTGCAAGATAACAGATTATCTCTGCTGGAATGTTGGGTATGGTGTCTGGA 1320
Db 1261 GTTTATGGCTCGAAGATAACAGATTATCTGCTCTGGAATGTTGGGCATGCTGTCTGGA 1320
QY 1321 CTTATTTCTGACTCCAGATCATCATCATCCACCAAGGAGACAGAACTGGATGCTGTAA 1380
Db 1321 CTTATTTCTGACTCCAGATCATCATCATCCACCAAGGAGACAGAACTGGATGCTGTAA 1380
QY 1381 AACATCCGGCTGGTAACAGTGCCTGCTGGCTGGGACATTCACCCGCGACCTCATTCCTAC 1440
Db 1381 AACATCCGGCTGGTAACAGTGCCTGCTGGCTGGGACATTCACCCGCGACCTCATTCCTAC 1440
QY 1441 ATCAATGAGTGGCTCCAAATAGACCTGGGGGAGGAGAAGATCGTGAAGGGCATCATCAAT 1500
Db 1441 ACCAATGAATGGCTCCAAAGTGGACCTGGGAGATGAGAAGATAGTAAGAGGTGTCTCATATT 1500
QY 1501 CAGGCTGGGAACCGGAGAGACAGGTGTTTCATGAGGAAGTTCAAGATCGGCTACAGC 1560
Db 1501 CAGGCTGGGAACCGGAGGAACAGGTGTTTCATGAGGAAGTTCAAGATCGGCTACAGT 1560
QY 1561 AACAAACGGCTCGACTGGAAGATGATCATGGATGACAGCAAAACGGAAGGCTCTTTT 1620
Db 1561 AACAAACGGCTCGACTGGAAGATGATCATGGATGACAGCAAAACGGAAGGCTCTTTT 1620
QY 1621 GAGGGAACAAACTATATATACATGAGCTGGGAGCTTTTCCAGCTCTCTCCACGGCA 1680
Db 1621 GAAGGGAACAAACTATATATGACAGCTGAGCTTCCGACGCTTTTCACTCTCTCCACAGG 1680
QY 1681 TTTCATCAGGATCTACCCGAGAGCCACTCATGCGGACTGCGGCTCAGATGGAGCTG 1740
Db 1681 TTTCATCAGGATCTACCCGAGAGCCACTCATGCGGACTGCGGCTCAGATGGAGCTG 1740
QY 1741 CTGGGCTGTGAAGTGAAGCCCTCAGCTGGACCGACACCTCCCAACGGGAACCTTGGTG 1800
Db 1741 CTGGGCTGTGAAGTGAAGCCCTCAGCTGGACCGACACCTCCCAACGGGAACCTTGGTG 1800
QY 1801 GATGATGTGATGACGACCGCCCACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
Db 1801 CATGATGTGATGACGACCGCCCACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
QY 1861 ACAGTGGGACCACTGTGCTGCCACAGAAAGCCCGCTCATGACAGCAGCACCATAACA 1920
Db 1861 ACAGGAGGACCACTGTGCTGCCACAGAAAGCCCGCTCATGACAGCAGCACCATAACA 1920
QY 1921 TCAGATTTTCCAACTATGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAGACTTTC 1980
Db 1921 TCAGATTTTCCCACTATGCTGTTTAACTGCGAGTTGGCTGGGGCTCTCACAAGACTTTC 1980
QY 1981 TGCCACTGGGAACATGACAACTACGCTGACGCTCAAGTGGAGTGTGTTGACCAAGAGG 2040
Db 1981 TGCCACTGGGAGCATGACAGCCATGACAGCTCAGTGGAGTGTGTTGACCAAGAGG 2040
QY 2041 GGACCCATTGAGGATCACAGAGATGCAACTTCATCTATTCCTCAAGCTGACGAAAT 2100
Db 2041 GGCGCGATTGAGGACCATACAGAGATGCAACTTCATCTATTCCTCAAGCTGATGAAAT 2100
QY 2101 CAGAAAGGCAAAAGTGGCTGCTGTGAGCCCTGTGGTTTATTCACAGAACTCTGCCAC 2160
Db 2101 CAGAAAGGCAAAAGTGGCTGCTGTGAGCCCTGTGGTTTATTCACAGAACTCTGCCAC 2160
QY 2161 TGCATGACCTTCTGGTATCACATGCTCTGGGTCGCCAGTCCGACACTCAGGCTCAAACTG 2220
Db 2161 TGTATGACCTTCTGGTATCACATGCTCTGGGTCGCCAGTCCGACACTCAGGCTCAAACTG 2220
QY 2221 CGCTACCAAGCCAGAGAGTACCATGAGTGGCTGATGGATGCCATTTGGACCAAGGT 2280
Db 2221 CGCTACCAAGCCAGAGAGTATGATCAACTGGCTGGATGGTGGTGGGACCAAGGA 2280
QY 2281 GACCACTGGAAGGAGGGCTCTTGTGCTCCACAGTCTCTGAACTTTATCAGGTATT 2340
Db 2281 GACCACTGGAAGAGGAGCTCTTGTGCTGCAAACTCTCTGAACTATATCAGGTATT 2340
QY 2341 TTCGAGGGCGAAATCGGAAAGGAAACCTTGGTGGGATTGCTGGATGACATTAGTATT 2400

Db 2341 TTTGAAGGTGAATCGGAAAGGAAACCTTGGTGGAAATTCGTGATCATCAGTATT 2400
QY 2401 AATAACCATTTTCAACAAGATTTCGCAAAACAGCAGACCTGGATAAAAGAACCCA 2460
Db 2401 AACAACCATTTTCTCAGGAAGACTGTGCAAAACCAACAGACCTAGATAAAAGAACACA 2460
QY 2461 GAAATTAATTTGATGAACAGGAGCGACGCCAGGATAGCAAGGTCAGGAGAGGTGAC 2520
Db 2461 GAAATTAATTTGATGAACAGGAGCGACCTCCAGGATATGAAGAGAGGAGGAGGTGAC 2520
QY 2521 AAGAATCTTCAGGAAGCCAGGCAATGTGTTGAAGACCTTGAACCCCTCCTCA 2575
Db 2521 AAGAATCTTCAGGAAGCCAGGCAATGTGTTGAATCTGCAAAATTCACCAACA 2575

RESULT 12
BC007533
LOCUS
DEFINITION Homo sapiens, neuropilin 1, clone MGC:15383 IMAGE:2958475, mRNA,
complete cds.
ACCESSION BC007533
VERSION BC007533.1 GI:14043096
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2232)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 22 Row: b Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomScan gene prediction.
FEATURES
Location/Qualifiers
1..2232
/organism="Homo sapiens"
/db_xref="LocusID:8829"
/db_xref="taxon:9606"
/clone="MGC:15383 IMAGE:2958475"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH_MGC_14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
261..2195
/codon_start=1
/product="neuropilin 1"
/protein_id="AAH07533.1"


```

/db_xref="GI:14043097"
/translation="MERGLPLLCALVALVLAAPAFNRNDKCGTIIKIESGYLTSFGY
PHSYHPSEKEMLIQADPDYQRTMINFNPHEDEDRCKYDYVEVDFGENENGHFRGK
FCCKIAPPVVSQPTFLIKFVSDYETHGAGFSIRVEIKRGPCEQSNYTPPSGVIKS
PGPEXPNSLECTYIVFAPKMSEIILFESFDELPDPPGGMFVCRDYLEIWDGPKM
DVGPHRYCGQTPKRRSSGILSKVFTYDLSAIKAGFSANVLSQSVSBDFKCM
EALUMSEGTESDITATSSOYNTWSAERSLAYPENGWTPGDSYFQWTFQVLDGLLR
FVFAVQTQASIKETKIKYVYTKIDVSNGEDWITIKEGNKPVLFQGNTPNTDVVV
AVPKPLTIRFVKIRATGTPFVSMRVEYCKITDYPCSKMLGMVSGLLSDQITSS
NOGRDNWPNIRLVTSRSGWALPPAPHYSINEWLOIDLGEKIVRGILIIQSGKHREN
KYPWRKFICYSNNGSDWKIMDDSKRKAASFEGNNNYDPELRTPPALSTRIRIYYP
TVLATEKPTVIDSTIOGSKIP*
BASE COUNT      621 a   546 c   558 g   507 t
ORIGIN

Query Match      69.3%; Score 1920.8; DB 9; Length 2232;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1922; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAGAGGGGGCTGCGCCTCTCTGGCGCGTGTGCGCCCTCGCTCTGCGCCCGCGGC 60
|||||
Db 261 ATGAGAGGGGGCTGCGCCTCTCTGGCGCGTGTGCGCCCTCGCTCTGCGCCCGCGGC 320
|||||

Qy 61 GCTTTTCGCAACGATGAATGTGGCGATACTATAAAAATTGAAAGCCCGGGTACCTTACA 120
|||||
Db 321 GCTTTTCGCAACGATAAATGTGGCGATACTATAAAAATTGAAAGCCCGGGTACCTTACA 380
|||||

Qy 121 TCTCCTGGTTATCCTCATCTTATACCCAAGTGA AAAATGCGAATGCGTGATTCAGGCT 180
|||||
Db 381 TCTCCTGGTTATCCTCATCTTATACCCAAGTGA AAAATGCGAATGCGTGATTCAGGCT 440
|||||

Qy 181 CCGGACCCATACAGAGAAATATGATCAACTCAACCTCACTTCGATTTGGAGGACAGA 240
|||||
Db 441 CCGGACCCATACAGAGAAATATGATCAACTCAACCTCACTTCGATTTGGAGGACAGA 500
|||||

Qy 241 GACTGCAAGTATGACTAGTGGAAAGTCTTCGATGGAGAAAATGAAAATGGACATTTTAGG 300
|||||
Db 501 GACTGCAAGTATGACTAGTGGAAAGTCTTCGATGGAGAAAATGAAAATGGACATTTTAGG 560
|||||

Qy 301 GGAAGTTCTGTGAAAGATAGCCCTCTCTCTGTTGTGCTTCAGGCGCAATTTCTTTT 360
|||||
Db 561 GGAAGTTCTGTGAAAGATAGCCCTCTCTCTGTTGTGCTTCAGGCGCAATTTCTTTT 620
|||||

Qy 361 ATCAAAATTTGCTCTGACTAGCAACACATGGTGACAGATTTTCCATACGTTATGA AATT 420
|||||
Db 621 ATCAAAATTTGCTCTGACTAGCAACACATGGTGACAGATTTTCCATACGTTATGA AATT 680
|||||

Qy 421 TTCAAGAGAGTCTCGAATGTTCCAGAACTACACACACCTAGTGGAGTGATAAAGTCC 480
|||||
Db 681 TTCAAGAGAGTCTCGAATGTTCCAGAACTACACACACCTAGTGGAGTGATAAAGTCC 740
|||||

Qy 481 CCCGGATTCCTGAAAATATFCCAACAGCCTTGAATGCACTTATATTTGCTTTTGC GCCA 540
|||||
Db 741 CCCGGATTCCTGAAAATATFCCAACAGCCTTGAATGCACTTATATTTGCTTTTGC GCCA 800
|||||

Qy 541 AAGATGTCAGAGATATTCCTGGAATTTGAAGCTTTTGCACCTGGAGCCTGACTCA AATCCT 600
|||||
Db 801 AAGATGTCAGAGATATTCCTGGAATTTGAAGCTTTTGCACCTGGAGCCTGACTCA AATCCT 860
|||||

Qy 601 CCAGGGGGGATGTTCTGTGCGTACGACCGGCTAGAAAATCTGGGATGGATTCCTTGATGTT 660
|||||
Db 861 CCAGGGGGGATGTTCTGTGCGTACGACCGGCTAGAAAATCTGGGATGGATTCCTTGATGTT 920
|||||

Qy 661 GGCCCTCACATTTGGCGGTTTACTGTGGACAGAAAACACACAGTTCGAATCCGATCCTCATCG 720
|||||
Db 921 GGCCCTCACATTTGGCGGTTTACTGTGGACAGAAAACACACAGTTCGAATCCGATCCTCATCG 980
|||||

Qy 721 GGCATTTCTCCATGGTTTTTTACACGACAGCCGATACCAAGAGAGTTTCTCAGCA 780
|||||
Db 981 GGCATTTCTCCATGGTTTTTTACACGACAGCCGATACCAAGAGAGTTTCTCAGCA 1040
|||||

Qy 781 AACTTACAGTGTCTTCAGAGACAGTGTCTCAGAAAGATTTCAAATGTATGGAAGCTTCTCGGC 840
|||||

```


QY 1681 TTCATCAGGATCTACCCCGAGAGAGACCTCATGCGGAGCTGGGGCTCAGAAATGGAGCTG 1740
|||||
Db 1967 TTCATCAGGATCTACCCCGAGAGAGACCTCATGCGGAGCTGGGGCTCAGAAATGGAGCTG 2026
QY 1741 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACCGACCTCCCAACCGGAACCTTGGTG 1800
|||||
Db 2027 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACCGACCTCCCAACCGGAACCTTGGTG 2086
QY 1801 GATGAATGTGATGACGACGACGACGCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
|||||
Db 2087 GATGAATGTGATGACGACGACGCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 2146
QY 1861 ACAGGTGGCACCCTGTGCTGGCCACAGAAAGCCACCGTCTATAGACAGCACCATAFCAA 1920
|||||
Db 2147 ACAGGTGGCACCCTGTGCTGGCCACAGAAAGCCACCGTCTATAGACAGCACCATAFCAA 2206
QY 1921 TCAG 1924
|||||
Db 2207 TCAG 2210
RESULT 14
AF145712
LOCUS AF145712 1772 bp mRNA linear PRI 21-MAR-2000
DEFINITION Homo sapiens soluble neuropilin-1 mRNA, complete cds.
ACCESSION AF145712
VERSION AF145712.1 GI:7271464
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1972)
AUTHORS Gagnon,M.L., Bielenberg,D.R., Gerechtman,Z., Miao,H.Q., Takashima,S.,
Soker,S. and Klagsbrun,M.
TITLE Identification of a natural soluble neuropilin-1 that binds
vascular endothelial growth factor: In vivo expression and
antitumor activity
Proc. Natl. Acad. Sci. U.S.A. 97 (6), 2573-2578 (2000)
JOURNAL 20183929
MEDLINE 10688880
PUBMED
REFERENCE 2 (bases 1 to 1972)
AUTHORS Gagnon,M.L. and Klagsbrun,M.
TITLE Direct Submission
JOURNAL Submitted (26-APR-1999) Surgical Research, Children's Hospital, 320
Longwood Ave., Boston, MA 02115, USA
FEATURES
source
1..1972
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10p12"
/cell_line="PC3"
/tissue_type="prostate adenocarcinoma"
1..1935
/function="receptor"
/note="SNRP1"
/codon_start=1
/product="soluble neuropilin-1"
/protein_id="AAFA4344.1"
/db_xref="GI:7271465"
/translation="MERGLPLLCVAVLALVLPAGAFNRDKGDTIKIESPGYLTSPGY
PHSHPSKCEWLQADPYQYIMINFPNFDLEDROCKDYVEVFGDENGNGHFRGK
FCGKTAPPVYSSGPFLLFKFVSDYETHGAGFTRYELFKRGPCSONYTPSGVIKS
PGFPEKPNSECTYIVAPKMSLILFEFSDLEPDSNPPGSMFCRDLRIEWDGFP
DVGPHRGICQKTPGRKSSGILSMVFYDFAKRGFSANTSVLOSSEDFKCM
BALGMESIEHSDQITASSQYSTNMSAERSLNYPNGWTPEGSYRERWIOVDLGLLR
FVTAVTGALISKETKKYVYKIDYSSNGEDWITIKEGNKPVLFOGNTNPDVLY
AVFPKPLTRVRIKPAWETGISMRFVYCKITDYGKMLGWSGLISDSOITSS
NOGDNMPENIRLVTSRSGWALPAPHYSINLWLOIDLGEKIVRGIIIOGKHREN
KYVFRKFKIGYSNNGSKMIMDSKRAKSAFEGNNYDTELTFFALSTRFRFIYP
ERATHGGGLRLMELGCEVEAPTAPGTPPNGLNVDCEDDQANCHSGTGDFDLTGTT
TVLATEKPTVIDSTIQSGIK"

BASE COUNT 562 a 479 c 481 g 450 t
ORIGIN
Query Match 69.2%; Score 1919.2; DB 9; Length 1972;
Best Local Similarity 99.8%; Pred. No. 0; Mismatches 3; Indels 0; Gaps 0;
Matches 1921; Conservative 0;
QY 1 ATGGAGAGGGGGTGGCGCTCTCTCGCGCGTCTCGCCCTCGTCTCGCCCGCGCGCG 60
|||||
Db 1 ATGGAGAGGGGGTGGCGCTCTCTCGCGCGTCTCGCCCTCGTCTCGCCCGCGCGCG 60
QY 61 GCTTTTCGCAAGATGAATGTGGCGATATATAAAATTTGAAAGCCCGGGTACCTTACA 120
|||||
Db 61 GCTTTTCGCAAGATGAATGTGGCGATATATAAAATTTGAAAGCCCGGGTACCTTACA 120
QY 121 TCTCTGGTTATCTCTCATTTATACCCCAAGTGAAGATGGAATGCGATGCTGATTCAGGCT 180
|||||
Db 121 TCTCTGGTTATCTCTCATTTATACCCCAAGTGAAGATGGAATGCGATGCTGATTCAGGCT 180
QY 181 CCGGACCCATACGAGAGAATATGATCAACTTCAACCCCTCACTTCGATTTGGAGGACAGA 240
|||||
Db 181 CCGGACCCATACGAGAGAATATGATCAACTTCAACCCCTCACTTCGATTTGGAGGACAGA 240
QY 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAAATGAAAATGGACATTTTAGG 300
|||||
Db 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAAATGAAAATGGACATTTTAGG 300
QY 301 GGAAAGTTCTGTGGAAAGATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
|||||
Db 301 GGAAAGTTCTGTGGAAAGATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
QY 361 ATCAAAATTTGTCTCTGACTTACGAAACACATGGTGCAGGATTTTTCATACGCTTATGAAAT 420
|||||
Db 361 ATCAAAATTTGTCTCTGACTTACGAAACACATGGTGCAGGATTTTTCATACGCTTATGAAAT 420
QY 421 TTCAGAGAGGTCCTGAATGTTTCCAGAACTACACAACTAGTGGAGTGTATGAAAGTCC 480
|||||
Db 421 TTCAGAGAGGTCCTGAATGTTTCCAGAACTACACAACTAGTGGAGTGTATGAAAGTCC 480
QY 481 CCGGATTCCTCGAAATATCCCAACAGCCTTGAATGACCTTATATCTCTCTCTCTCTCTCTCT 540
|||||
Db 481 CCGGATTCCTCGAAATATCCCAACAGCCTTGAATGACCTTATATCTCTCTCTCTCTCTCTCT 540
QY 541 AAGATGTCAGAGATATCTCGAATTTGAAAGCTTTGACCTGGAGCTGACTCAATCTCT 600
|||||
Db 541 AAGATGTCAGAGATATCTCGAATTTGAAAGCTTTGACCTGGAGCTGACTCAATCTCT 600
QY 601 CCAGGGGGATGTTCTGCTGCTACGACCGGTAGAAAATCTGGGATGGATTCCTCTGATGTT 660
|||||
Db 601 CCAGGGGGATGTTCTGCTGCTACGACCGGTAGAAAATCTGGGATGGATTCCTCTGATGTT 660
QY 661 GGCCCTCACATTTGGGGTGTCTGTGACGAGAAAACACAGGTCGAAATCCGATCCCTCATCG 720
|||||
Db 661 GGCCCTCACATTTGGGGTGTCTGTGACGAGAAAACACAGGTCGAAATCCGATCCCTCATCG 720
QY 721 GGCAATTTCTCCATGGTTTTTATACCCGACGAGCGGATAGCAAAAAGAGGTTTCTCAGCA 780
|||||
Db 721 GGCAATTTCTCCATGGTTTTTATACCCGACGAGCGGATAGCAAAAAGAGGTTTCTCAGCA 780
QY 781 AACTACAGTGTCTTGAGAGAGTGTCTCAGAAAGATTTCAAATGTATGGAAGCTCTGGGC 840
|||||
Db 781 AACTACAGTGTCTTGAGAGAGTGTCTCAGAAAGATTTCAAATGTATGGAAGCTCTGGGC 840
QY 841 ATGGAATCAGGAAATTCATTCTGACCATGATACAGCTTCTTCCAGTATAGCAACCAAC 900
|||||
Db 841 ATGGAATCAGGAAATTCATTCTGACCATGATACAGCTTCTTCCAGTATAGCAACCAAC 900
QY 901 TGGTCTGACAGCGCTCCCGCTGAACCTGAGATGGGTGGAGTCCCGGAGAGGAT 960
|||||
Db 901 TGGTCTGACAGCGCTCCCGCTGAACCTGAGATGGGTGGAGTCCCGGAGAGGAT 960
QY 961 TCTTACCAGAGTGGATACAGTATGGGCTTCTGCGCTTCTGTCACAGGCTCTCGGG 1020
|||||

Db 961 TCCTACCGAGAGTGGATACAGGTAGACTTGGGCCCTTCTCGCTTTGTACCGGCTGTCGGG 1020
Qy 1021 ACACAGGGCGGCATTTCAAAGAAAACCAAGAGAAATATTATGTCAAGACTTTACAAGATC 1080
Db 1021 ACACAGGGCGGCATTTCAAAGAAAACCAAGAGAAATATTATGTCAAGACTTTACAAGATC 1080
Qy 1081 GACGTTAGCTCCACGGGGAAGACTGGATCACCAATAAAGAGGAACAAACCTGTTCTC 1140
Db 1081 GACGTTAGCTCCACGGGGAAGACTGGATCACCAATAAAGAGGAACAAACCTGTTCTC 1140
Qy 1141 TTTCAGGGAACACCAACCCACAGATGTTGGTTCAGATATTCCTCCCAACCACTGATA 1200
Db 1141 TTTCAGGGAACACCAACCCACAGATGTTGGTTCAGATATTCCTCCCAACCACTGATA 1200
Qy 1201 ACTCGATTGTCGGAATCAAGCTCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1260
Db 1201 ACTCGATTGTCGGAATCAAGCTCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1260
Qy 1261 GTATACGGTTGCAAGATAACAGATTATCTTCTGCTTGAATGTTGGGTATGTTGCTGGA 1320
Db 1261 GTATACGGTTGCAAGATAACAGATTATCTTCTGCTTGAATGTTGGGTATGTTGCTGGA 1320
Qy 1321 CTTATTTCTGACTCCAGATCACATCATCCAAACCAAGAGACAGAACTGGATGCTGAA 1380
Db 1321 CTTATTTCTGACTCCAGATCACATCATCCAAACCAAGAGACAGAACTGGATGCTGAA 1380
Qy 1381 AACATCCGCTGGTAACAGTCGCTGCTGGTGGGACATTCACCCGACCTCATTCCTAC 1440
Db 1381 AACATCCGCTGGTAACAGTCGCTGCTGGTGGGACATTCACCCGACCTCATTCCTAC 1440
Qy 1441 ATCAATGATGTGCTCCAAATAGACCTGGGGAGGAGAGATCGTGGGGGATCATCAT 1500
Db 1441 ATCAATGATGTGCTCCAAATAGACCTGGGGAGGAGAGATCGTGGGGGATCATCAT 1500
Qy 1501 CAGGTGGGAACACCGGAGAGAACAAAGGTGTTTCATGAGGAAGTTCAAGATCGGCTAC 1560
Db 1501 CAGGTGGGAACACCGGAGAGAACAAAGGTGTTTCATGAGGAAGTTCAAGATCGGCTAC 1560
Qy 1561 AACACGCTCGGACTGGAAGATGATCATGGATGACACCAACGCAAGCGGAAGTCTTTT 1620
Db 1561 AACACGCTCGGACTGGAAGATGATCATGGATGACACCAACGCAAGCGGAAGTCTTTT 1620
Qy 1621 GAGGCAACAACTATGATACACTGAGCTGCGGACTTTTCCAGCTCTCTCCACGGGA 1680
Db 1621 GAGGCAACAACTATGATACACTGAGCTGCGGACTTTTCCAGCTCTCTCCACGGGA 1680
Qy 1681 TTCATCAGATCTACCCGAGAGAGCCACTCATGCGGACTGGGCTCAGAAATGAGCTG 1740
Db 1681 TTCATCAGATCTACCCGAGAGAGCCACTCATGCGGACTGGGCTCAGAAATGAGCTG 1740
Qy 1741 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACCGACCACTCCCAACGGGAACCTGGT 1800
Db 1741 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACCGACCACTCCCAACGGGAACCTGGT 1800
Qy 1801 GATGAATGTGATGACGACGACCACTGCCACAGTGAACAGGTGATGACTTCCAGCTC 1860
Db 1801 GATGAATGTGATGACGACGACCACTGCCACAGTGAACAGGTGATGACTTCCAGCTC 1860
Qy 1861 ACAGTGGCAGCAGCTGTGCTGGCCACAGAAAAGCCCGCTCATAGACAGACCAATACAA 1920
Db 1861 ACAGTGGCAGCAGCTGTGCTGGCCACAGAAAAGCCCGCTCATAGACAGACCAATACAA 1920
Qy 1921 TCAG 1924
Db 1921 TCAG 1924

RESULT 15
LOCUS AF280547 3750 bp mRNA linear PRI 20-DEC-2000
DEFINITION Homo sapiens neurophilin-1 soluble isoform 11 (NRPI), complete
cds, alternatively spliced.
ACCESSION AF280547

AF280547.1 GI:11907931
VERSION 1
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3750)
AUTHORS Rossignol, M., Gagnon, M.L. and Klagsbrun, M.
TITLE Genomic organization of human neurophilin-1 and neurophilin-2 genes: identification and distribution of splice variants and soluble isoforms
JOURNAL Genomics 70 (2), 211-222 (2000)
MEDLINE 20564205
PUBMED 11112349
REFERENCE 2 (bases 1 to 3750)
AUTHORS Rossignol, M., Gagnon, M.L. and Klagsbrun, M.
TITLE Direct Submission
JOURNAL Submitted (21-JUN-2000) Surgical Research, Children's Hospital, 320 Ave, Boston, MA 02115, USA
FEATURES
Location/Qualifiers
1..3750
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10p12"
/cell_line="PC3"
/tissue_type="prostate adenocarcinoma"
1..3750
/gene="NRPI"
1..2115
/gene="NRPI"
/note="receptor for VEGF165 and semaphorins class 3;
SLINRPI; generated by alternative pre-mRNA processing in
intron 11 of NRPI"
/codon_start=1
/product="neurophilin-1 soluble isoform 11"
/protein_id="AAG41406.1"
/db_xref="GI:11907932"
/translation="MERGLPLCALVALVAPAGAFRNDKGDITIKIESPCYLTSPGV
PHSYHPSEKCEWLIQADPDYQRIINFNPHFDLEDRCYDYVEYFGEENHNGHFRK
FCGKIAPPVVSQGFLLIFKPSDIETHGAGFSIRYIFRKGPECSONYTRTPSIVIKS
PGPEKYPNSLECTIVFAPKMSIILEFSDPSNPPGFCRDRLEIWDGFP
DVGFPHIGYCGQKTPGRISRSGLSNVFTDSAIAKEGFSANTSVLQSSVSEDFCK
EALCMSEITHSDQITASSQSYSTNWSERLNYPENGWTPGDSYREWIQVDLGLLR
FVAVGTQGAISKETKKYVKYKIDVSSNGEDWITIKENKPLVPGQNTNPDVVV
AVFKPLITREVRKIPATWETGISMREVEYCKITDPCSGMLGMSGLISDSITSS
NQDRNMPENIRLVTSGWALPPAPHSITINWLIQDLEEKIVRGIIQGGKHREN
KVFMRFKIGYSNNGSDMKIMDDSKRKAISFEKNNDYDPELRTFFPALSTRFRIYIP
ERATHGLGLRMELLGCEVAPTAGPTTPNGNLVDECDQDQANCHSGTGDQFQLTGAE
TIFPLLYHFSSCLSWDLTPVCVLVTPHGRELPNRNSCLARTASSFPFHVWIDELF
LIATTICNNLSHPESORLGLS"
BASE COUNT 1108 a 841 c 806 g 995 t
ORIGIN
Query Match 67.2%; Score 1862.2; DB 9; Length 3750;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1864; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGGAGAGGGGCTGCGCTCTCTGCGCGCTGCTGCGCCCTCGTCTCGCCGCCCGCCG 60
Db 1 ATGGAGAGGGGCTGCGCTCTCTGCGCGCTGCTGCGCCCTCGTCTCGCCGCCCGCCG 60
Qy 61 GCTTTTCCCAAGGATGTGGCGATACTATAAAATGAAGCCCGGCTACCTTACA 120
Db 61 GCTTTTCCCAAGGATGTGGCGATACTATAAAATGAAGCCCGGCTACCTTACA 120
Qy 121 TCTCTGTTATCTCTATCTTATCACCAAGTGAATAATGCGATGCTGATTCAGCT 180
Db 121 TCTCTGTTATCTCTATCTTATCACCAAGTGAATAATGCGATGCTGATTCAGCT 180
Qy 181 CCGGACCCATACCAAGAGAAATATGATCAACTTCAACCTCTACTTCGATTTGAGGACAGA 240
Db 181 CCGGACCCATACCAAGAGAAATATGATCAACTTCAACCTCTACTTCGATTTGAGGACAGA 240

QY 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAAAATGAAATGGCAATTTAGG 300
Db 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAAAATGAAATGGCAATTTAGG 300
QY 301 GGAAGCTCTGTGGAAGATAGCCCTCTCTCTGTTGTGTCTTCAGGGCCATTTCTTTT 360
Db 301 GGAAGCTCTGTGGAAGATAGCCCTCTCTCTGTTGTGTCTTCAGGGCCATTTCTTTT 360
QY 361 ATCAAAATTTGCTCTGACTACGAAACACATGGTGCAGGATTTTCCATACGTTATGAAAT 420
Db 361 ATCAAAATTTGCTCTGACTACGAAACACATGGTGCAGGATTTTCCATACGTTATGAAAT 420
QY 421 TTCAGAGAGGTCTGAAATGTTCCAGAACTACACAACTAGTGGAGTGATAAAGTCC 480
Db 421 TTCAGAGAGGTCTGAAATGTTCCAGAACTACACAACTAGTGGAGTGATAAAGTCC 480
QY 481 CCGGATTCCTTGAAATATCCACAGCCTTGATGACATTATATTTCTTTGGCCA 540
Db 481 CCGGATTCCTTGAAATATCCACAGCCTTGATGACATTATATTTCTTTGGCCA 540
QY 541 AAGATCTCAGAGATTATCCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAAATCCT 600
Db 541 AAGATCTCAGAGATTATCCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAAATCCT 600
QY 601 CAGGGGGATGTTCTGCTGCTACGACCGGCTAGAAATCTGGATGGATTTCCCTGATGTT 660
Db 601 CAGGGGGATGTTCTGCTGCTACGACCGGCTAGAAATCTGGATGGATTTCCCTGATGTT 660
QY 661 GGCCCTCACATTTGGGGTGTACTGTGACAGAAACACAGGTGCAATCCGATCCCTCATCG 720
Db 661 GGCCCTCACATTTGGGGTGTACTGTGACAGAAACACAGGTGCAATCCGATCCCTCATCG 720
QY 721 GGCATCTCTCCATGTTTTTACACCGACAGCGGATAGCAAAAGAGTTTCTCAGCA 780
Db 721 GGCATCTCTCCATGTTTTTACACCGACAGCGGATAGCAAAAGAGTTTCTCAGCA 780
QY 781 AACTACAGTGTCTGAGAGAGTGTCTCAGAGATTTTCAAATGTATGGAAGCTTGGGC 840
Db 781 AACTACAGTGTCTGAGAGAGTGTCTCAGAGATTTTCAAATGTATGGAAGCTTGGGC 840
QY 841 ATGGAATCAGGAGAAATTCATCTGACAGATCACAGCTTCTTCCAGTATAGCACCAAC 900
Db 841 ATGGAATCAGGAGAAATTCATCTGACAGATCACAGCTTCTTCCAGTATAGCACCAAC 900
QY 901 TGGTCTGCAGAGCGCTCCCGCTGAACTACCTGAGAATGGGTGGACTCCCGGAGAGGAT 960
Db 901 TGGTCTGCAGAGCGCTCCCGCTGAACTACCTGAGAATGGGTGGACTCCCGGAGAGGAT 960
QY 961 TCCACCGAGTGGATACAGGTAGACTTGGGCTTCTCGGCTTTGTACAGGCTGTGCGG 1020
Db 961 TCCACCGAGTGGATACAGGTAGACTTGGGCTTCTCGGCTTTGTACAGGCTGTGCGG 1020
QY 1021 ACACAGGCGCCATTTCAAAGAAACCAAGAGAAATATATGTCAGAGCTTACAAGATC 1080
Db 1021 ACACAGGCGCCATTTCAAAGAAACCAAGAGAAATATATGTCAGAGCTTACAAGATC 1080
QY 1081 GAGCTTAGCTCCAAAGGGAAGCTGGATCACCATAAAGAGGAAACAAACCTGTTCTC 1140
Db 1081 GAGCTTAGCTCCAAAGGGAAGCTGGATCACCATAAAGAGGAAACAAACCTGTTCTC 1140
QY 1141 TTTACGGGAAACACCAACCCACAGATGTTGTTGCTAGTATCCCCAAACCACTGATA 1200
Db 1141 TTTACGGGAAACACCAACCCACAGATGTTGTTGCTAGTATCCCCAAACCACTGATA 1200
QY 1201 ACTCGATTTGTCGATCAGGCTGCACTTGGGAACTGGGCAATATCTATGAGATTTGAA 1260
Db 1201 ACTCGATTTGTCGATCAGGCTGCACTTGGGAACTGGGCAATATCTATGAGATTTGAA 1260
QY 1261 GTATACGGTTGCAAGATACAGATTATCTTCTCTGGAATGTTGGGTATGTTGTTGGA 1320
Db 1261 GTATACGGTTGCAAGATACAGATTATCTTCTCTGGAATGTTGGGTATGTTGTTGGA 1320

QY 1321 CTTATTTCTGACTCCAGATCACATCATCAACCAAGGAGACAGAAACTGGATGCTGAA 1380
Db 1321 CTTATTTCTGACTCCAGATCACATCATCAACCAAGGAGACAGAAACTGGATGCTGAA 1380
QY 1381 AACATCCGCTGGTAACCACTGCTCTGGCTGGCACTTCCACCCGACCTCATTTCTTAC 1440
Db 1381 AACATCCGCTGGTAACCACTGCTCTGGCTGGCACTTCCACCCGACCTCATTTCTTAC 1440
QY 1441 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAAGATCGTAGGGGCATCATATT 1500
Db 1441 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAAGATCGTAGGGGCATCATATT 1500
QY 1501 CAGGGTGGGAAGACCGAGAGAACAAAGTGTTCATGAGGAAGTTCAAGATCGGGTACAGC 1560
Db 1501 CAGGGTGGGAAGACCGAGAGAACAAAGTGTTCATGAGGAAGTTCAAGATCGGGTACAGC 1560
QY 1561 AACACGGCTCGGACTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Db 1561 AACACGGCTCGGACTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
QY 1621 GAGGCAACAACAACATATGATACACCTGAGCTCGGACTTTTCCAGCTCTCTCCACGCA 1680
Db 1621 GAGGCAACAACAACATATGATACACCTGAGCTCGGACTTTTCCAGCTCTCTCCACGCA 1680
QY 1681 TTATCAGGATCTACCCCGAGAGACCACTCATGCGGACTGGGGCTCAGATGAGGCTG 1740
Db 1681 TTATCAGGATCTACCCCGAGAGACCACTCATGCGGACTGGGGCTCAGATGAGGCTG 1740
QY 1741 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACGACCACTCCCAACGGGAACTTGGTG 1800
Db 1741 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACGACCACTCCCAACGGGAACTTGGTG 1800
QY 1801 GATGAATGTGATGACGACCAAGCCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
Db 1801 GATGAATGTGATGACGACCAAGCCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
QY 1861 ACAGGTG 1867
Db 1861 ACAGGTG 1867

Search completed: November 19, 2002, 23:16:59
Job time : 4844 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	2772	100.0	2772	20	AAx08993	Human semaphorin r
2	2772	100.0	2772	21	AAZ31461	Human neuropilin c
3	2762.4	99.7	5653	20	AAx59151	Human neuropilin-1
4	2762.4	99.7	5653	20	AAx81726	DNA encoding human
5	2751.2	99.2	2772	20	AAx79349	Human VEGF(165) R/N
6	2743.6	99.0	4146	22	AAH99711	Human protein enco
7	2741.4	98.9	2775	22	AAF83309	Human neuropilin-1
8	2536.6	91.5	2568	22	AAF83311	Human NP-1 extracel
9	2258	81.5	2808	22	AAx06049	Angiotensin conver

PA (REGC) UNIV CALIFORNIA.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted and is derived by analysis of the total score distribution.

PI Chen H, He Z, Tessier-lavigne M;
XX WPI; 1999-120780/10.
DR P-PSDB; AAW96246.
XX
PT New semaphorin receptor (SR) class 1 and class 2 proteins - useful
XX for, e.g. regulating nerve cell function and morphology
XX
PS Claim 3; Page 41-42; 105pp; English.
XX
CC Semaphorin receptor (SR) proteins are involved in nerve cell
CC guidance. The SR proteins can regulate cell function and morphology,
CC especially nerve cells. Inhibitors of SR proteins promote nerve cell
CC growth. Binding agents specific to the SR polypeptides are useful
CC for diagnosis and therapy, especially where disease or disease
CC prognosis is associated with improper or undesirable axon outgrowth
CC or orientation. The SR nucleic acid sequences are used to provide
CC translatable transcripts, hybridisation probes, PCR primers,
CC diagnostic nucleic acids, etc. They are used in detecting the
CC presence of SR genes, and to identify wild type and mutant alleles.
CC Mutant alleles are used to generate allele-specific oligonucleotide
CC probes for clinical diagnosis. The methods are used to identify
CC agents which modulate the activity of SR polypeptides. This
CC sequence encodes the human SR1 polypeptide.
XX
SQ Sequence 2772 BP; 772 A; 664 C; 702 G; 634 T; 0 other;

Query Match 100.0%; Score 2772; DB 20; Length 2772;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2772; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGAGGGGTCGCGCTCTCTCGCGCGTCTCGCCCTCGCTCGCCCGCGCGGC 60
DB 1 ATGGAGAGGGGTCGCGCTCTCTCGCGCGTCTCGCCCTCGCTCGCCCGCGCGGC 60

QY 61 GCTTTTCGAACGATGAATGTGGCGTACTATAAAATTTGAAGCCCGGGTACCTTACA 120
DB 61 GCTTTTCGAACGATGAATGTGGCGTACTATAAAATTTGAAGCCCGGGTACCTTACA 120

QY 121 TCTCCTGGTTATCCTCATTTATACCCCAAGTGAATAATGGAATGGCTGATTCAGGCT 180
DB 121 TCTCCTGGTTATCCTCATTTATACCCCAAGTGAATAATGGAATGGCTGATTCAGGCT 180

QY 181 CCGGACCCATACAGAGAATTATGATCAACTTCAACCCCTCACTTCGATTGGAGGACAGA 240
DB 181 CCGGACCCATACAGAGAATTATGATCAACTTCAACCCCTCACTTCGATTGGAGGACAGA 240

QY 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAAATGAAAATGGACATTTAGG 300
DB 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAAATGAAAATGGACATTTAGG 300

QY 301 GGAAAGTTCTGTGGAAGATAGCCCTCCTCTGTGTGTCTTTCAGGGCCATTTCTTTT 360
DB 301 GGAAAGTTCTGTGGAAGATAGCCCTCCTCTGTGTGTCTTTCAGGGCCATTTCTTTT 360

QY 361 ATCAAAATTTGTCTGACTACGAAACACATGTTGTCAGGATTTTCCATACGTTATGAAAT 420
DB 361 ATCAAAATTTGTCTGACTACGAAACACATGTTGTCAGGATTTTCCATACGTTATGAAAT 420

QY 421 TTCAAGAGAGGTCCTGAATGTTCCAGAACTACACACACCTAGTGGAGTGATAAGTCC 480
DB 421 TTCAAGAGAGGTCCTGAATGTTCCAGAACTACACACACCTAGTGGAGTGATAAGTCC 480

QY 481 CCGGGATTCCTTGAAATATCCCAACAGCCTTGAATGACATTTATATTGCTTTGGGCCA 540
DB 481 CCGGGATTCCTTGAAATATCCCAACAGCCTTGAATGACATTTATATTGCTTTGGGCCA 540

QY 541 AAGATGTCAGAGATTATCTGGAATTTGAAAGCTTTTGACCTGGAGCCTGACTCAATCCT 600
DB 541 AAGATGTCAGAGATTATCTGGAATTTGAAAGCTTTTGACCTGGAGCCTGACTCAATCCT 600

QY 601 CCAGGGGGATGTTCTGCTGCTACGACCGGCTAGAAATCTGGATGGATTCCTGATGTT 660
DB 601 CCAGGGGGATGTTCTGCTGCTACGACCGGCTAGAAATCTGGATGGATTCCTGATGTT 660

DB 601 CCAGGGGGGATGTTCTGCTGCTACGACCGGCTAGAAATCTGGGATGGATTCCTGATGTT 660
QY 661 GGCCCTCACATTTGGGCGTTACTGTGGACAGAAAACACACAGGTGGAATCCGATCCTCATCG 720
DB 661 GGCCCTCACATTTGGGCGTTACTGTGGACAGAAAACACACAGGTGGAATCCGATCCTCATCG 720
QY 721 GGCAATTCCTCCATGTTTATACCCGACAGCGCATAGCAAAAGAGTTTCTCAGCA 780
DB 721 GGCAATTCCTCCATGTTTATACCCGACAGCGCATAGCAAAAGAGTTTCTCAGCA 780
QY 781 AACTACAGTGTCTTCAGAGCAGTGTCTCAGAGATTTCAAAATGTATGGAGCTCTGGGC 840
DB 781 AACTACAGTGTCTTCAGAGCAGTGTCTCAGAGATTTCAAAATGTATGGAGCTCTGGGC 840
QY 841 ATGGAATCAGGAGAAATTCATTTCTGACCCAGATCACAGCTTCTTCCAGTATAGCAACAAC 900
DB 841 ATGGAATCAGGAGAAATTCATTTCTGACCCAGATCACAGCTTCTTCCAGTATAGCAACAAC 900
QY 901 TGGTCTGACAGAGCGCTCCCGCTGAACTACCCCTGAGAATGGGTGGACTCCCGGAGAGGAT 960
DB 901 TGGTCTGACAGAGCGCTCCCGCTGAACTACCCCTGAGAATGGGTGGACTCCCGGAGAGGAT 960
QY 961 TCCCTACGAGAGTGGATACAGGTAGACTTGGGCTTCTGCGCTTGTGACGGGCTGCGGG 1020
DB 961 TCCCTACGAGAGTGGATACAGGTAGACTTGGGCTTCTGCGCTTGTGACGGGCTGCGGG 1020
QY 1021 ACACAGGGGCCATTTCAAAGAAACCAAGAGAAATATATGCTCAAGACTTACAAGATC 1080
DB 1021 ACACAGGGGCCATTTCAAAGAAACCAAGAGAAATATATGCTCAAGACTTACAAGATC 1080
QY 1081 GAGCTTAGCTTCCAAACGGGGAAGACTGGATCACCATAAAGAGAAACAAACCTGTTCTC 1140
DB 1081 GAGCTTAGCTTCCAAACGGGGAAGACTGGATCACCATAAAGAGAAACAAACCTGTTCTC 1140
QY 1141 TTTTCAGGGAACACCAACCCACAGATGTTGTTGTCAGATATTCCTCCAAACCACTGATA 1200
DB 1141 TTTTCAGGGAACACCAACCCACAGATGTTGTTGTCAGATATTCCTCCAAACCACTGATA 1200
QY 1201 ACTCGATTTGTCGGAATCAAGCCTGCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1260
DB 1201 ACTCGATTTGTCGGAATCAAGCCTGCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1260
QY 1261 GTATACGGTTGCAAGATAACAGATTTATCTGCTCTGGAATGTTGGGTATGTTGTTCTGGA 1320
DB 1261 GTATACGGTTGCAAGATAACAGATTTATCTGCTCTGGAATGTTGGGTATGTTGTTCTGGA 1320
QY 1321 CTATTTTCTGACTCCAGATCACATCATCAACCAAGGAGACAGAACTGGATGCTGAA 1380
DB 1321 CTATTTTCTGACTCCAGATCACATCATCAACCAAGGAGACAGAACTGGATGCTGAA 1380
QY 1381 AACATCCGCTGTGAACCACTGCTCTGCTGGGCACTTCCACCCGCACTCATTTCTCTAC 1440
DB 1381 AACATCCGCTGTGAACCACTGCTCTGCTGGGCACTTCCACCCGCACTCATTTCTCTAC 1440
QY 1441 ATCAATGATGGCTCCAAATAGACCTTGGGGGAGGAGAGATCTGTAGGGGCACTCATATT 1500
DB 1441 ATCAATGATGGCTCCAAATAGACCTTGGGGGAGGAGAGATCTGTAGGGGCACTCATATT 1500
QY 1501 CAGGGTGGGAAGCACCAGAGACAGAGTGTTCATCAGGAAGTTCAAGATTCGGGTACAGC 1560
DB 1501 CAGGGTGGGAAGCACCAGAGACAGAGTGTTCATCAGGAAGTTCAAGATTCGGGTACAGC 1560
QY 1561 AACACGGCTCGGACTTGGGAAGATGATCATGGATGACAGCAAAACGCAAGCGAAGTCTTTT 1620
DB 1561 AACACGGCTCGGACTTGGGAAGATGATCATGGATGACAGCAAAACGCAAGCGAAGTCTTTT 1620
QY 1621 GAGGGCAACAACAATATGATACACTGAGCTGCGGACTTTTCCAGCTCTCTCCACGCGA 1680
DB 1621 GAGGGCAACAACAATATGATACACTGAGCTGCGGACTTTTCCAGCTCTCTCTCCACGCGA 1680
QY 1681 TTTCATCAGGATCTACCCCGAGAGAGCCACTCATGCGGACTGGGGCTCAGATGGAGCTG 1740
DB 1681 TTTCATCAGGATCTACCCCGAGAGAGCCACTCATGCGGACTGGGGCTCAGATGGAGCTG 1740

241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAAAATGAAAATGGACATTTTAGG 300
241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAAAATGAAAATGGACATTTTAGG 300
301 GGAAGTTCTGTGGAAAGATAGCCCTCTCTCTGTGTGTCTTCAGGGCCATTTCTTTT 360
301 GGAAGTTCTGTGGAAAGATAGCCCTCTCTCTGTGTGTCTTCAGGGCCATTTCTTTT 360
361 ATCAAAATTTGTCTGTGACTACGAAACACATGTTGCGAGGATTTTCCATACGTTATGAAAT 420
361 ATCAAAATTTGTCTGTGACTACGAAACACATGTTGCGAGGATTTTCCATACGTTATGAAAT 420
421 TTCAAGAGAGGTCCTGAAATGTTTCCAGAACTACACAACTAGTGGAGTGATAAGTCC 480
421 TTCAAGAGAGGTCCTGAAATGTTTCCAGAACTACACAACTAGTGGAGTGATAAGTCC 480
481 CCGGATTCCTTGAAATATCCCAACAGCCTTGAATGCACATATATGTTCTTGGGCCA 540
481 CCGGATTCCTTGAAATATCCCAACAGCCTTGAATGCACATATATGTTCTTGGGCCA 540
541 AAGATGTCAGAGATTAATCTGGAATTTGAAAGCTTTGACCTGGAGCTGACCTCAAAATCCT 600
541 AAGATGTCAGAGATTAATCTGGAATTTGAAAGCTTTGACCTGGAGCTGACCTCAAAATCCT 600
601 CCAGGGGGATGTTCTGTCGTACGACCGGCTAGAAATCTGGGATGGATTCCTCTGATGTT 660
601 CCAGGGGGATGTTCTGTCGTACGACCGGCTAGAAATCTGGGATGGATTCCTCTGATGTT 660
661 GGGCTCAGATTTGGGGTACTGTGGACAGAAACACAGGTCGAATCCGATCCTCATCG 720
661 GGGCTCAGATTTGGGGTACTGTGGACAGAAACACAGGTCGAATCCGATCCTCATCG 720
721 GCATTCCTCTCCATGTTTACACCGACAGCGGATAGCAAAAGATTTCTCAGCA 780
721 GGCATTCCTCTCCATGTTTACACCGACAGCGGATAGCAAAAGATTTCTCAGCA 780
781 AACTACAGTGTCTGAGAGCAGTGTCTCAGAAGATTTTCAAATGTATGGAAGCTCTGGGC 840
781 AACTACAGTGTCTGAGAGCAGTGTCTCAGAAGATTTTCAAATGTATGGAAGCTCTGGGC 840
841 ATGGATCAGGAGAAATTCATTTGACACAGATCAGAGCTTCTCCAGTATAGCACCAC 900
841 ATGGATCAGGAGAAATTCATTTGACACAGATCAGAGCTTCTCCAGTATAGCACCAC 900
901 TGGTCTGACAGAGCTCCCGCTGAACTACCTGAGAATGGGTGGACTCCCGAGAGGAT 960
901 TGGTCTGACAGAGCTCCCGCTGAACTACCTGAGAATGGGTGGACTCCCGAGAGGAT 960
961 TCTACCGAGAGTGGATACAGGTAGACTTGGGCTTCTCGCTTTGTACAGGCTGTCTGGG 1020
961 TCTACCGAGAGTGGATACAGGTAGACTTGGGCTTCTCGCTTTGTACAGGCTGTCTGGG 1020
1021 ACACAGGGCCATTTCAAAGAAACCAAGAGAAATATATGTCAGAGCTTACAGATC 1080
1021 ACACAGGGCCATTTCAAAGAAACCAAGAGAAATATATGTCAGAGCTTACAGATC 1080
1081 GAGCTTAGCTTCAACGGGAAAGACTGGATCACCATAAAGAGAGAAACAACTGTCTC 1140
1081 GAGCTTAGCTTCAACGGGAAAGACTGGATCACCATAAAGAGAGAAACAACTGTCTC 1140
1141 TTTCAGGGAACACCAACCCACAGATGTTGTTGGTTCAGTATTCCTCCAAACCACTGATA 1200
1141 TTTCAGGGAACACCAACCCACAGATGTTGTTGGTTCAGTATTCCTCCAAACCACTGATA 1200
1201 ACTCGATTTGTCGAATCAAGCTGCACTTGGGAACCTGGCATATCTATGAGATTTGAA 1260
1201 ACTCGATTTGTCGAATCAAGCTGCACTTGGGAACCTGGCATATCTATGAGATTTGAA 1260
1261 GTATACGGTTGCAAGATAACAGATATCTCTGCTCTGGAATGTTGGGATGGTGTCTGA 1320
1261 GTATACGGTTGCAAGATAACAGATATCTCTGCTCTGGAATGTTGGGATGGTGTCTGA 1320
1321 CTATTTCTGACTCCAGATACATCATCCAAACCAAGAGACAGAAACTGGATGCTGAA 1380

1321 CTATTTCTGACTCCAGATACATCATCAACCAAGGAGACAGAACTGGATGCTGAA 1380
1381 AACATCCGCTCGTAAACAGTCTGCTGCTGGCACTTCCACCCGACCTCATTTCTCTAC 1440
1381 AACATCCGCTCGTAAACAGTCTGCTGCTGGCACTTCCACCCGACCTCATTTCTCTAC 1440
1441 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAAATCGTGAGGGGCATCATCAT 1500
1441 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAAATCGTGAGGGGCATCATCAT 1500
1501 CAGGGTGGAGACCGAGAGAACAAAGTGTTCATGAGGAAGTTCAAGATCGGGTACAGC 1560
1501 CAGGGTGGAGACCGAGAGAACAAAGTGTTCATGAGGAAGTTCAAGATCGGGTACAGC 1560
1561 AACACGGCTCGGACTGGAAGATGATCATGGATGACAGCAAAACGCAAGCGAAGTCTTTT 1620
1561 AACACGGCTCGGACTGGAAGATGATCATGGATGACAGCAAAACGCAAGCGAAGTCTTTT 1620
1621 GAGGGCAACAACAACATATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCACGCA 1680
1621 GAGGGCAACAACAACATATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCACGCA 1680
1681 TTCAATCAGGATCTACCCCGAGAGACCACTCATGCGGGACTGGGGCTCAGATGGAGCTG 1740
1681 TTCAATCAGGATCTACCCCGAGAGACCACTCATGCGGGACTGGGGCTCAGATGGAGCTG 1740
1741 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACCCACCACTCCCAACGGGAACCTTTGGT 1800
1741 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACCCACCACTCCCAACGGGAACCTTTGGT 1800
1801 GATGAATGTGATGACGACCAAGCCAACTGCCAGTGGAGCAAGTGTGATGACTTCCAGCTC 1860
1801 GATGAATGTGATGACGACCAAGCCAACTGCCAGTGGAGCAAGTGTGATGACTTCCAGCTC 1860
1861 ACAGTGGCACCACTGTGCTGGCCACAGAAAGCCCAAGGTCATAGACAGCACCATTACAA 1920
1861 ACAGTGGCACCACTGTGCTGGCCACAGAAAGCCCAAGGTCATAGACAGCACCATTACAA 1920
1921 TCAGAGTTTCCAACTATGTTTAACTGTGAATTTGGTGGGCTCTCACAAGACTTTC 1980
1921 TCAGAGTTTCCAACTATGTTTAACTGTGAATTTGGTGGGCTCTCACAAGACTTTC 1980
1981 TGCCACTGGGAACATGACAACTCACAGTGCAGCTCAAGTGGAGTGTGTTGACCAAGAGAG 2040
1981 TGCCACTGGGAACATGACAACTCACAGTGCAGCTCAAGTGGAGTGTGTTGACCAAGAGAG 2040
2041 GGACCCATTCAGGATCACACAGGAGATGSCAACTTCTATTTCCCAAGCTGACGAAAT 2100
2041 GGACCCATTCAGGATCACACAGGAGATGSCAACTTCTATTTCCCAAGCTGACGAAAT 2100
2101 CAGAAGGGCAAGTGGCTCGCTGGTGGAGCCCTGTTTATTTCCCAAGACTCTGCCAC 2160
2101 CAGAAGGGCAAGTGGCTCGCTGGTGGAGCCCTGTTTATTTCCCAAGACTCTGCCAC 2160
2161 TGCTAGACCTTCTGSPATCACATGCTGGGTCCCAGTCCGACACTCAGGGTCAAAGT 2220
2161 TGCTAGACCTTCTGSPATCACATGCTGGGTCCCAGTCCGACACTCAGGGTCAAAGT 2220
2221 CGGTACCAAGCCAGAGAGTACGATCAGCTGGTGTGGATGGCCATTTGGACACCAAGGT 2280
2221 CGGTACCAAGCCAGAGAGTACGATCAGCTGGTGTGGATGGCCATTTGGACACCAAGGT 2280
2281 GACCACTGGAAGGAGGGCTGTCTTGTCTCCAACTCTCTGAACTTTATCAGGTGAT 2340
2281 GACCACTGGAAGGAGGGCTGTCTTGTCTCCAACTCTCTGAACTTTATCAGGTGAT 2340
2341 TTTCAGGGCGAAATCGGAAAAGAAACCTTGGTGGGATTTGCTGTGGATGACATTAAT 2400
2341 TTTCAGGGCGAAATCGGAAAAGAAACCTTGGTGGGATTTGCTGTGGATGACATTAAT 2400
2401 AATAACCACTTTTCACAGAAAGATTTGTGCAAAACACAGACCTGATGATAAAAGACCCA 2460
2401 AATAACCACTTTTCACAGAAAGATTTGTGCAAAACACAGACCTGATGATAAAAGACCCA 2460

Db 2401 AATAACCACATTTCACAAGAAGATTGTGCAAAACACAGACAGCTGGATAAAAAGAACCCCA 2460
QY 2461 GAAATTAATAATTGATGAACAGGAGCAGCCAGGATAGCAAGGTGAAGGAGAAGGTGAC 2520
Db 2461 GAAATTAATAATTGATGAACAGGAGCAGCCAGGATAGCAAGGTGAAGGAGAAGGTGAC 2520
QY 2521 AGAACATCTCAGGAAGCCAGCAATGTGTTGAAGACCTTGAACCCATCCTCATCACCC 2580
Db 2521 AGAACATCTCAGGAAGCCAGCAATGTGTTGAAGACCTTGAACCCATCCTCATCACCC 2580
QY 2581 ATCATAGCCATGAGCGCCCTGGGGTCCCTGGGGGCTGTGTGGGCTCGTGTGAC 2640
Db 2581 ATCATAGCCATGAGCGCCCTGGGGTCCCTGGGGGCTGTGTGGGCTCGTGTGAC 2640
QY 2641 TGTGCTGTGTGGCATAATGGGATGTGAGAAAGAAACTTGTCTGCCCTGGAGAACTATAAC 2700
Db 2641 TGTGCTGTGTGGCATAATGGGATGTGAGAAAGAAACTTGTCTGCCCTGGAGAACTATAAC 2700
QY 2701 TTTGAACCTTGTGATGCTGAAGTTGAAAAAGACAAACTGAATACACAGAGTACTTAT 2760
Db 2701 TTTGAACCTTGTGATGCTGAAGTTGAAAAAGACAAACTGAATACACAGAGTACTTAT 2760
QY 2761 TCGGAGGCATGA 2772
Db 2761 TCGGAGGCATGA 2772

RESULT 3
AAX59151
ID AAX59151 standard; cDNA; 5653 BP.
XX
AC AAX59151;
XX
DT 06-SEP-1999 (first entry)
XX
DE Human neuropilin-1 cDNA.
XX
KW Neuropilin-1; human; vascular endothelial growth factor receptor;
KW VEGF165R; inhibitor; metastasis; angiogenesis; inflammation;
KW arthritis; diabetic retinopathy; cytostatic; antiinflammatory;
KW antiangiogenic; antidiabetic; ophthalmological; antiarthritic; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 239..3010
FT /tag= a
FT sig_peptide 239..301
FT /tag= b
FT mat_peptide 302..3007
FT /tag= c
XX
PN W09929858-A1.
XX
PD 17-JUN-1999.
XX
PF 09-DEC-1998; 98WO-US26138.
XX
PR 09-SEP-1998; 98US-0099615.
PR 09-DEC-1997; 97US-0069155.
PR 12-DEC-1997; 97US-0069687.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Gagnon ML, Klagsbrun M, Soker S;
XX
DR WPI; 1999-394975/33.
DR P-PSDB; AAY06317.
XX
PT Soluble human neuropilins and related polynucleotides
XX
PS Example 1; Page 81-82; 98pp; English.

CC This is the nucleotide sequence of cDNA coding for human
CC neuropilin-1 (see AAY06317), or VEGF165R, a novel receptor of
CC vascular endothelial growth factor (VEGF). Expression of
CC neuropilin is associated with angiogenesis and motile metastatic
CC cancer cells, and thus is an important target for antiangiogenic
CC and anticancer therapy. Neuropilin-1 cDNA was obtained from
CC MDA-MB-231 cells by expression cloning. A soluble neuropilin (see
CC AAY06319) comprising amino acids 227-587 of the full-length
CC neuropilin-1 is claimed, and can be isolated from neuropilin
CC producing cells or is recombinantly engineered from neuropilin
CC encoding DNA. Soluble truncated neuropilins are capable of
CC binding to VEGF165 or any form of VEGF that contains exon 7 and are
CC therefore useful for inhibiting VEGF interaction not only with
CC neuropilins but also with KDR/Flk-1 and Flt1. The soluble NPs may
CC also act as dominant negative receptors when expressed in cells by
CC dimerizing with intact neuropilin receptors. The soluble
CC neuropilins are useful for treatment of diseases or disorders
CC associated with VEGF, e.g. metastasis, inappropriate angiogenesis,
CC chronic inflammation, diabetic retinopathy and arthritis. Solid
CC tumors can also be treated. Polynucleotides encoding neuropilins
CC can be used in the preparation of medicaments for treating diseases
CC or disorders associated with VEGF.
XX
SQ Sequence 5653 BP; 1614 A; 1202 C; 1271 G; 1566 T; 0 other;

Query Match 99.7%; Score 2762.4; DB 20; Length 5653;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2766; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGGAGAGGGGCTGCCGCTCTCTCGCGCGTGTGCTGCTCGCCGGCGGCG 60
Db 239 ATGGAGAGGGGCTGCCGCTCTCTCGCGCGTGTGCTGCTCGCCGGCGGCG 298
QY 61 GCTTTTCGCAACGATGAATGTGCGGATACTATAAAATTAAGAGCCCGGTACCTTACA 120
Db 299 GCTTTTCGCAACGATGAATGTGCGGATACTATAAAATTAAGAGCCCGGTACCTTACA 358
QY 121 TCTCTGTTATCTCTCATCTTATCACCAAGTGAAAAATGCGAATGGCTGATTCAGGCT 180
Db 359 TCTCTGTTATCTCTCATCTTATCACCAAGTGAAAAATGCGAATGGCTGATTCAGGCT 418
QY 181 CCGGACCCATACAGAGAAATATGATCAACTCAACCCCTCACTTCGATTTGGAGGACAGA 240
Db 419 CCGGACCCATACAGAGAAATATGATCAACTCAACCCCTCACTTCGATTTGGAGGACAGA 478
QY 241 GACTGCAAGTATGACTACCTGGAAGTCTTCGATGGAGAAATCAAAATGGACATTTAGG 300
Db 479 GACTGCAAGTATGACTACCTGGAAGTCTTCGATGGAGAAATCAAAATGGACATTTAGG 538
QY 301 GGAAAGTTCTGTGGAAAGATAGCCCTCCTCTGTGTGTCTTCAGGGGCATTTCTTTT 360
Db 539 GGAAAGTTCTGTGGAAAGATAGCCCTCCTCTGTGTGTCTTCAGGGGCATTTCTTTT 598
QY 361 ATCAAAATTTGCTCTGACTACGAAACACATGGTGCAGGATTTTCCATAGTATGAAAT 420
Db 599 ATCAAAATTTGCTCTGACTACGAAACACATGGTGCAGGATTTTCCATAGTATGAAAT 658
QY 421 TTCAAGAGAGGTCCTGAATGTTCCCAAGAACTACACACACCTAGTGGAGTATAAGTCC 480
Db 659 TTCAAGAGAGGTCCTGAATGTTCCCAAGAACTACACACACCTAGTGGAGTATAAGTCC 718
QY 481 CCGGATTCCTTGAAAAATATCCAAACAGCTTGAATGCACCTATATTTCTTTGGCCCA 540
Db 719 CCGGATTCCTTGAAAAATATCCAAACAGCTTGAATGCACCTATATTTCTTTGGCCCA 778
QY 541 AAGATGTCAGAGATTAATCTCGAAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAACTCT 600
Db 779 AAGATGTCAGAGATTAATCTCGAAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAACTCT 838
QY 601 CCAGGGGGGATGTTCTGTGCTACGACCGGCTAGAAATCTGGGATGGATTCCTGTATGTT 660
Db 839 CCAGGGGGGATGTTCTGTGCTACGACCGGCTAGAAATCTGGGATGGATTCCTGTATGTT 998

QY 661 GGCCCTCACATTTGGCGCTTACTGTGGACAGAAACACACAGGTGCAATTCGATCCTCATCG 720
DB 899 GGCCCTCACATTTGGCGCTTACTGTGGACAGAAACACACAGGTGCAATTCGATCCTCATCG 958
QY 721 GGCAATCTCTCCATGGTTTTTACACGACAGCGCGATAGCAAAAGAGGTTTTCTCAGCA 780
DB 959 GGCAATCTCTCCATGGTTTTTACACGACAGCGCGATAGCAAAAGAGGTTTTCTCAGCA 1018
QY 781 AACTACAGTGTCTTGACAGACAGTGTCTCAGAAGATTTCAAATGTATGGAAGTCTCTGGC 840
DB 1019 AACTACAGTGTCTTGACAGACAGTGTCTCAGAAGATTTCAAATGTATGGAAGTCTCTGGC 1078
QY 841 ATGGAATCAGGAGAAATTCATTTGACACAGATCACAGCTTTCTCCCAAGTATAGCACCAAC 900
DB 1079 ATGGAATCAGGAGAAATTCATTTGACACAGATCACAGCTTTCTCCCAAGTATAGCACCAAC 1138
QY 901 TGGTCTGCAGAGCGCTCCCGCTTGAACCTACCTTGAGATCGGTGGACTCCCGAGAGGAT 960
DB 1139 TGGTCTGCAGAGCGCTCCCGCTTGAACCTACCTTGAGATCGGTGGACTCCCGAGAGGAT 1198
QY 961 TCCTACCGAGAGTGGATACAGGTAGACTTGGGCTTCTGCGCTTTGTGCACGGCTGTGGG 1020
DB 1199 TCCTACCGAGAGTGGATACAGGTAGACTTGGGCTTCTGCGCTTTGTGCACGGCTGTGGG 1258
QY 1021 ACACAGGGCGCCATTTCAAAGAAACCAAGAGAAATATTTATGTCAAGACTTACAAGATC 1080
DB 1259 ACACAGGGCGCCATTTCAAAGAAACCAAGAGAAATATTTATGTCAAGACTTACAAGATC 1318
QY 1081 GAGCTTAGCTCCACCGGGAGACTGGATCACCATAAAGAAAGAAACAACCTGTCTC 1140
DB 1319 GAGCTTAGCTCCACCGGGAGACTGGATCACCATAAAGAAAGAAACAACCTGTCTC 1378
QY 1141 TTTTCAGGGAACACCAACCCACAGATGTTGTGTTGTCAGTATTTCCCAACCACTGATA 1200
DB 1379 TTTTCAGGGAACACCAACCCACAGATGTTGTGTTGTCAGTATTTCCCAACCACTGATA 1438
QY 1201 ACTCGATTTCTCGAATCAAGCCTGCACTTTGGGAACTGGCATACTATGAGATTTGAA 1260
DB 1439 ACTCGATTTCTCGAATCAAGCCTGCACTTTGGGAACTGGCATACTATGAGATTTGAA 1498
QY 1261 GTATACGTTTGAAGATAACAGATATTCCTTGCTCTGGAATGTTGGTATGTTGCTGGA 1320
DB 1499 GTATACGTTTGAAGATAACAGATATTCCTTGCTCTGGAATGTTGGTATGTTGCTGGA 1558
QY 1321 CTTATTTCTGACTCCAGATCACATCATCAACCAAGGAGACAGAACTGGATGCCGTGAA 1380
DB 1559 CTTATTTCTGACTCCAGATCACATCATCAACCAAGGAGACAGAACTGGATGCCGTGAA 1618
QY 1381 AACATCGGCTGTAAACCACTGCTCTGGCTGGGACCTTCCACCCGACCTCATTCCTAC 1440
DB 1619 AACATCGGCTGTAAACCACTGCTCTGGCTGGGACCTTCCACCCGACCTCATTCCTAC 1678
QY 1441 ATCAATGAGTGGCTCCAAATFAGACTGGGGAGGAGAAAGATCGTGAGGGGCATCATCAT 1500
DB 1679 ATCAATGAGTGGCTCCAAATFAGACTGGGGAGGAGAAAGATCGTGAGGGGCATCATCAT 1738
QY 1501 CAGGGTGGGAAGCACCGAGAGAACAGGTGTTTCATGAGGAAGTTCAAGATCGGGTACAGC 1560
DB 1739 CAGGGTGGGAAGCACCGAGAGAACAGGTGTTTCATGAGGAAGTTCAAGATCGGGTACAGC 1798
QY 1561 AACACGGCTCGGACTGGGAAGATGATCGATGAGCAGCAAAACCAAGGCGAAGTCTTTT 1620
DB 1799 AACACGGCTCGGACTGGGAAGATGATCGATGAGCAGCAAAACCAAGGCGAAGTCTTTT 1858
QY 1621 GAGGGCAACAACACTATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCACGCGA 1680
DB 1859 GAGGGCAACAACACTATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCACGCGA 1918
QY 1681 TTTTCATCAGGATCTACCCCGAGAGACCACTCATGCGGGACTGGGGCTCAGATGGAGCTG 1740
DB 1919 TTTTCATCAGGATCTACCCCGAGAGACCACTCATGCGGGACTGGGGCTCAGATGGAGCTG 1978
QY 1741 CTGGGCTGTGAAGTGGAAAGCCCTTACAGCTGGACCGACCTCCCAACGGGAACCTTGGTG 1800

DB 1979 CTGGGCTGTGAAGTGGAAAGCCCTTACAGCTGGACCGACCACTCCCAACGGGAACCTTGGTG 2038
QY 1801 GATGAATGTGATGACGACGACGAGCCCACTGCCACAGTGGACAGAGTGATGACTTCCAGCTC 1860
DB 2039 GATGAATGTGATGACGACGACGAGCCCACTGCCACAGTGGACAGAGTGATGACTTCCAGCTC 2098
QY 1861 ACAGGTGGCACCACTGTGCTGGCCACAGAAAAGCCACCGTCTATAGACAGCACCATAACAA 1920
DB 2099 ACAGGTGGCACCACTGTGCTGGCCACAGAAAAGCCACCGTCTATAGACAGCACCATAACAA 2158
QY 1921 TCAGAGTTTCCAAATATATGTTTTTAACTGTGAATTTTGGCTGGGCTCTCAAGACCTTC 1980
DB 2159 TCAGAGTTTCCAAATATATGTTTTTAACTGTGAATTTTGGCTGGGCTCTCAAGACCTTC 2218
QY 1981 TGGCCTTGGGAAACATGACATCAGTGCACCTCAAGTGGAGTGTGTGACCAAGCAAGACG 2040
DB 2219 TGGCCTTGGGAAACATGACATCAGTGCACCTCAAGTGGAGTGTGTGACCAAGCAAGACG 2278
QY 2041 GGACCCATTCAGGATCACACAGGAGATGGCAACTTTCATCTATTTCCCAAGCTGACGAAAAT 2100
DB 2279 GGACCCATTCAGGATCACACAGGAGATGGCAACTTTCATCTATTTCCCAAGCTGACGAAAAT 2338
QY 2101 CAGAAGGGCAAACTGGCTCGCCTGGTGAGCCCTGTGTTTTTATTTCCCAAGACTCTGCCAC 2160
DB 2339 CAGAAGGGCAAACTGGCTCGCCTGGTGAGCCCTGTGTTTTTATTTCCCAAGACTCTGCCAC 2398
QY 2161 TGCATGACCTTCTGGTATCACATGCTGSGTCCCACTGCGGACACCTCAGGTCACAACTG 2220
DB 2399 TGCATGACCTTCTGGTATCACATGCTGSGTCCCACTGCGGACACCTCAGGTCACAACTG 2458
QY 2221 CGCTACCAAGAGCCAGAGGAGTACGATCAGCTGCTGGATGGCCATTTGACACCAAGAGT 2280
DB 2459 CGCTACCAAGAGCCAGAGGAGTACGATCAGCTGCTGGATGGCCATTTGACACCAAGAGT 2518
QY 2281 GACCACTGGAAGGAAGGCGCTGCTTCTGCCACAAAGTCTCTGAACTTTATCAGGTGATT 2340
DB 2519 GACCACTGGAAGGAAGGCGCTGCTTCTGCCACAAAGTCTCTGAACTTTATCAGGTGATT 2578
QY 2341 TTCAGSGGGCAAACTCGSAAAGGAAACCTTGGTGGGATGCTGTGGATGACATTAGTATT 2400
DB 2579 TTCAGSGGGCAAACTCGSAAAGGAAACCTTGGTGGGATGCTGTGGATGACATTAGTATT 2638
QY 2401 AATAACCACTTTCACAAGAAGATTGTGCAAAACACAGCAGACCTGGGATAAAAAAGACCCA 2460
DB 2639 AATAACCACTTTCACAAGAAGATTGTGCAAAACACAGCAGACCTGGGATAAAAAAGACCCA 2698
QY 2461 GAAATTTAAATTTGATGAAACAGGAGGACGCCAGGATACGAAGGTGAAGGAGAGGTGAC 2520
DB 2699 GAAATTTAAATTTGATGAAACAGGAGGACGCCAGGATACGAAGGTGAAGGAGAGGTGAC 2758
QY 2521 AAGAACATCTCCAGGAGGCCAGGCAATGTGTTGAAGACCTTAGAACCTTCCTCATCACC 2580
DB 2759 AAGAACATCTCCAGGAGGCCAGGCAATGTGTTGAAGACCTTAGAACCTTCCTCATCACC 2818
QY 2581 ATCATAGCCATGAGCCCTGGGGCTCTCTGGGGCTGTCTGTGGGGTGTGCTGTGCTGTAC 2640
DB 2819 ATCATAGCCATGAGTGCCTTGGGGTCTCTCTGGGGCTGTCTGTGGGGTGTGCTGTGCTGTAC 2878
QY 2641 TGTGCTGTGGCATTAATGGGATGTCAGAAAGAACTTGTCTGCCCTGGGAGACTATAAC 2700
DB 2879 TGTGCTGTGGCATTAATGGGATGTCAGAAAGAACTTGTCTGCCCTGGGAGACTATAAC 2938
QY 2701 TTTTGAATCTGGATGCTGTGAAGTTGAAAAAGACAACTGTAATACACAGACTACTTAT 2760
DB 2939 TTTTGAATCTGGATGCTGTGAAGTTGAAAAAGACAACTGTAATACACAGACTACTTAT 2998
QY 2761 TCGAGGGCATGA 2772
DB 2999 TCGAGGGCATGA 3010

RESULT 4

Db 1559 CTTATTTCAGATCCAGATCACATCATCAACCAAGGGAGCAGAAACTGGATGCCGTGAA 1618
QY 1381 AACATCCGCCTGGTAACCAAGTCGCTCGCTGGGCACTTCCACCCGACCTCATTCCTTAC 1440
Db 1619 AACATCCGCCTGGTAACCAAGTCGCTCGCTGGGCACTTCCACCCGACCTCATTCCTTAC 1678
QY 1441 ATCAATAGTGGCTCCAAATAGACCTCGGGGAGGAGAAGATCGTAGGGGCACTCATCATTT 1500
Db 1679 ATCAATAGTGGCTCCAAATAGACCTCGGGGAGGAGAAGATCGTAGGGGCACTCATCATTT 1738
QY 1501 CAGGCTGGGAAGCACCAGAGACAGAGTGTTCATCAGGAAGTTCAAGATCGGGTACAGC 1560
Db 1739 CAGGCTGGGAAGCACCAGAGACAGAGTGTTCATCAGGAAGTTCAAGATCGGGTACAGC 1798
QY 1561 AACACCGCTCGGACTGGAAGATGATCATGGATGACAGCAAAACGCAAGCGGAAGTCTTTT 1620
Db 1799 AACACCGCTCGGACTGGAAGATGATCATGGATGACAGCAAAACGCAAGCGGAAGTCTTTT 1858
QY 1621 GAGGCAACAACAATATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCAGCGGA 1680
Db 1859 GAGGCAACAACAATATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCAGCGGA 1918
QY 1681 TTCAATCAGATCTACCCGAGAGAGCCACTCATGCGGGACTGGGGCTCAGAATGGAGCTG 1740
Db 1919 TTCAATCAGATCTACCCGAGAGAGCCACTCATGCGGGACTGGGGCTCAGAATGGAGCTG 1978
QY 1741 CTGGGCTGTGAAGTGGAAAGCCCTACAGCTGGACCGACCACTCCCAACGGGAAGTGGTG 1800
Db 1979 CTGGGCTGTGAAGTGGAAAGCCCTACAGCTGGACCGACCACTCCCAACGGGAAGTGGTG 2038
QY 1801 GATGAATGTGATGACAGCAGCGCAACTGCCAGTGGGAACAGGTGATGACTTCCAGCTC 1860
Db 2039 GATGAATGTGATGACAGCAGCGCAACTGCCAGTGGGAACAGGTGATGACTTCCAGCTC 2098
QY 1861 ACAGGTGGCACCACCTGTCTGCCACAGAAAGCCACCGTATAGACAGCACCATACAA 1920
Db 2099 ACAGGTGGCACCACCTGTCTGCCACAGAAAGCCACCGTATAGACAGCACCATACAA 2158
QY 1921 TCAGAGTTTCCAAACATATGGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAGACTTTC 1980
Db 2159 TCAGAGTTTCCAAACATATGGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAGACTTTC 2218
QY 1981 TGGCACTGGGAACATGACAATCACGTGACGCTCAAGTGGAGTGTGTGACCAAGAGCG 2040
Db 2219 TGGCACTGGGAACATGACAATCACGTGACGCTCAAGTGGAGTGTGTGACCAAGAGCG 2278
QY 2041 GGACCCATTTCAGATCACACAGGAGATGCAACTTCATCTATTTCCCAAGCTGACGNAAT 2100
Db 2279 GGACCCATTTCAGATCACACAGGAGATGCAACTTCATCTATTTCCCAAGCTGACGNAAT 2338
QY 2101 CAGAAGGGCAAGTGGCTCGCTGGTGGAGCCCTGTGGTTTATTTCCAGAACTCTGCCAC 2160
Db 2339 CAGAAGGGCAAGTGGCTCGCTGGTGGAGCCCTGTGGTTTATTTCCAGAACTCTGCCAC 2398
QY 2161 TGCATGACCTTCGTGATCACATGTCTGGGTCCCGCTCGGACACTTCAGGCTCAAACTG 2220
Db 2399 TGCATGACCTTCGTGATCACATGTCTGGGTCCCGCTCGGACACTTCAGGCTCAAACTG 2458
QY 2221 CGGTACCAAGCCAGAGAGTACGATACGTGGTCTGGATGGCCATTTGGACACCAAGT 2280
Db 2459 CGGTACCAAGCCAGAGAGTACGATACGTGGTCTGGATGGCCATTTGGACACCAAGT 2518
QY 2281 GACCACTGGAAGGAGGGGCTGTCTGCTCCCAAGTCTCTGAACTTTATCAGGTGATT 2340
Db 2519 GACCACTGGAAGGAGGGGCTGTCTGCTCCCAAGTCTCTGAACTTTATCAGGTGATT 2578
QY 2341 TTCAGGGCGAAATCGGAAAAGAAACCTTGGTGGATTCGTGGATGACATTAAGTATT 2400
Db 2579 TTCAGGGCGAAATCGGAAAAGAAACCTTGGTGGATTCGTGGATGACATTAAGTATT 2638
QY 2401 AATAACCAATTTCAAGAAAGATTTGCAAAACCCAGCAGACCTGGATAAAAAGAACCA 2460

Db 2639 AACACCAACATTTTCAAGAAGATTGTGCAAAACCCAGCAGACCTGGATAAAAAGAACCCA 2698
QY 2461 GAAATTTAAATTTGATGAAACAGGAGCAGCCAGGATAGCAAGGTGAAGGAGAAGGTGAC 2520
Db 2699 GAAATTTAAATTTGATGAAACAGGAGCAGCCAGGATACGNAAGTGAAGGAGAAGGTGAC 2758
QY 2521 AGAACATCTCCAGGAAGCCAGGCAATGTGTGAAGACCTTAGAACCCCATCTCATCACC 2580
Db 2759 AAGAACATCTCCAGGAAGCCAGGCAATGTGTGAAGACCTTAGAACCCCATCTCATCACC 2818
QY 2581 ATCATAGCCATGAGCCCTCGGGGTCCTCTGGGGCTGTCTGTGGGTCGTGCTGTAC 2640
Db 2819 ATCATAGCCATGAGTCCTCGGGGTCCTCTGGGGCTGTCTGTGGGTCGTGCTGTAC 2878
QY 2641 TGTGCTGTGTGCATATATGGGATGTGAGAAAGAAACTTCTGCCCCTGGAGAACTATAAC 2700
Db 2879 TGTGCTGTGTGCATATATGGGATGTGAGAAAGAAACTTCTGCCCCTGGAGAACTATAAC 2938
QY 2701 TTTGAACCTTGTGGATGTGAAAGTTGAAAGAAACTGAATACACAGAGTACTTAT 2760
Db 2939 TTTGAACCTTGTGGATGTGAAAGTTGAAAGAAACTGAATACACAGAGTACTTAT 2998
QY 2761 TCGGAGGCATGA 2772
Db 2999 TCGGAGGCATGA 3010
RESULT 5
AAx79349
ID AAx79349 standard; DNA; 2772 BP.
XX
AC AAx79349;
DT
DT 14-SEP-1999 (first entry)
XX
DE Human VEGF(165)R/NP-1 gene.
XX
KW Human; vascular endothelial growth factor; VEGF; receptor; neuropilin;
KW antagonist; semaphorin; collapsin; cancer; metastasis; angiogenesis;
KW diabetic retinopathy; osteoarthritis; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO929729-A2.
XX
PD 17-JUN-1999.
XX
PF 09-DEC-1998; 98WO-US26114.
PR 19-MAR-1998; 98US-0078541.
PR 09-DEC-1997; 97US-0069155.
PR 12-DEC-1997; 97US-0069687.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Klagsbrun M, Miao H, Soker S, Takashima S;
DR WPI: 1999-394959/33.
DR P-PSDB; AAY14562.
XX
PT Neuropilin receptor antagonists is useful for treatment of, e.g.
PT metastatic cancer
XX
PS Disclosure; Fig 14; 80pp; English.
XX
CC This sequence represents the coding sequence for the human VEGF(165)R/NP1
CC (or neuropilin-1). The invention relates to a neuropilin antagonist that
CC binds neuropilin and has VEGF antagonistic activity as determined by the
CC human umbilical vein endothelial cell (HUVEC) proliferation assay using
CC VEGF165. The neuropilin antagonist (semaphorin/collapsin family member)
CC is useful for treatment of a disease or disorder associated with VEGF,
CC including cancer such as metastatic cancer and angiogenesis diseases,
CC diabetic retinopathy, osteoarthritis and psoriasis.

Dd 527 GACTCGAAGTATGACTACGTGGAGTCTTCGATGGAGAAATGAAATGGACATTTTAGG 586
Qy 301 GGAAGATTCGTGGAAAGATAGCCCTCTCTCTGTGTGTCTTCAGGGCCATTTCTTTTT 360
Dd 587 GGAAGATTCGTGGAAAGATAGCCCTCTCTCTGTGTGTCTTCAGGGCCATTTCTTTTT 646
Qy 361 ATCAAATTTGCTCTGACTACGAAACACATGGTCAGGATTTTCCATACGTTATGAAAT 420
Dd 647 ATCAAATTTGCTCTGACTACGAAACACATGGTCAGGATTTTCCATACGTTATGAAAT 706
Qy 421 TTCAAGAGAGTCCCTGAATGTTCCAGAACTACACACACCTAGTGGAGTGAATAAGTCC 480
Dd 707 TTCAAGAGAGTCCCTGAATGTTCCAGAACTACACACACCTAGTGGAGTGAATAAGTCC 766
Qy 481 CCCGATTCCTCGAAAAATATCCCAACAGCCTTGAATGCACATATA-TTGTCTTTGGCCG 539
Dd 767 CCCGATTCCTCGAAAAATATCCCAACAGCCTTGAATGCACATATA-TTGTCTTTGGCCG 826
Qy 540 AAAGATGTACAGAGATTATCCT- GGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAATC 598
Dd 827 AAAGATGTACAGAGATTATCCTTGGATTTTGAAGCTTTGACCTGGAGCCTGACTCAAATC 886
Qy 599 CTCAGGGGGGATGTTCTGCTACGACCGCTAGAAATCTGGGATGGATTCCTGTATG 658
Dd 887 CTCAGGGGGGATGTTCTGCTACGACCGCTAGAAATCTGGGATGGATTCCTGTATG 946
Qy 659 TTGGCCCTCACATTTGGCGTTACTGTGGACAGAAACACCAAGCTCGAATCCGATCCTCAT 718
Dd 947 TTGGCCCTCACATTTGGCGTTACTGTGGACAGAAACACCAAGCTCGAATCCGATCCTCAT 1006
Qy 719 CGGGCATTCCTCCATGTTTTTTACCCGACAGCGCGATAGCAAAAGAGTTTTCTCAG 778
Dd 1007 CGGGCATTCCTCCATGTTTTTTACCCGACAGCGCGATAGCAAAAGAGTTTTCTCAG 1066
Qy 779 CAACTACAGTGTCTGCAGAGCACTGTCTCAGAGATTTCAATGTATGGAGCTCTGG 838
Dd 1067 CAACTACAGTGTCTGCAGAGCACTGTCTCAGAGATTTCAATGTATGGAGCTCTGG 1126
Qy 839 GCATGGAATCAGGAGAAATTCATTTGACCAGATCACAGCTTCTTCCAGTATACACCA 898
Dd 1127 GCATGGAATCAGGAGAAATTCATTTGACCAGATCACAGCTTCTTCCAGTATACACCA 1186
Qy 899 ACTGCTGTGCAGAGCGCTCCCGCTGAACTACCTTGAGATGGGTGGACTCCCGGAGAG 958
Dd 1187 ACTGCTGTGCAGAGCGCTCCCGCTGAACTACCTTGAGATGGGTGGACTCCCGGAGAG 1246
Qy 959 ATTCTCTACGAGATGATACAGTGTAGACTTGGCGCTTCTGCGCTTGTCTCAGCGCTGCG 1018
Dd 1247 ATTCTCTACGAGATGATACAGTGTAGACTTGGCGCTTCTGCGCTTGTCTCAGCGCTGCG 1306
Qy 1019 GGACACAGGGGCCATTTCAAAGAAACCAAGAAATATTTATGTCAAGACTTACAAGA 1078
Dd 1307 GGACACAGGGGCCATTTCAAAGAAACCAAGAAATATTTATGTCAAGACTTACAAGA 1366
Qy 1079 TCGAGGTTAGCTTCCAAACGGGGAAGACTGGATCACCATAAAAGAGAAACAAACCTGTC 1138
Dd 1367 TCGAGGTTAGCTTCCAAACGGGGAAGACTGGATCACCATAAAAGAGAAACAAACCTGTC 1426
Qy 1139 TCTTTTCAGGAAACACCAACCCACAGATGTTGTGGTTCAGTATTTCCCAACCACTGA 1198
Dd 1427 TCTTTTCAGGAAACACCAACCCACAGATGTTGTGGTTCAGTATTTCCCAACCACTGA 1486
Qy 1199 TAACTCGATTTCTCGAATCAAGCTCGAACTTGGAACTGGCAATCTATGAGATTG 1258
Dd 1487 TAACTCGATTTCTCGAATCAAGCTCGAACTTGGAACTGGCAATCTATGAGATTG 1546
Qy 1259 AAGTATACGGTTGCAAGATAACAGATTATCCTTGTCTTGGATTTGGGTATGTTGCTG 1318
Dd 1547 AAGTATACGGTTGCAAGATAACAGATTATCCTTGTCTTGGATTTGGGTATGTTGCTG 1606
Qy 1319 GACTTATTTCTGACTCCAGATCACATCATCAACCAAGGAGACAGAACTGGATGCTG 1378
Dd 1607 GACTTATTTCTGACTCCAGATCACATCATCAACCAAGGAGACAGAACTGGATGCTG 1666

Qy 1379 AAAACATCGCCTGTTAAACCAGTCTGCTGGCTGGGCACTTCCACCCGACCTCATTTCT 1438
Dd 1667 AAAACATCGCCTGTTAAACCAGTCTGCTGGCTGGGCACTTCCACCCGACCTCATTTCT 1726
Qy 1439 ACATCAATAGTGGCTCCAAATAGACCTTGGGGAGGAGAAGATCTGTAGGGGCATCATCA 1498
Dd 1727 ACATCAATAGTGGCTCCAAATAGACCTTGGGGAGGAGAAGATCTGTAGGGGCATCATCA 1786
Qy 1499 TTCAGGGTGGNAGCACCGAGAGAACAGGTGTTTCATGAGGAAGTTCAAGATCGGGTACA 1558
Dd 1787 TTCAGGGTGGNAGCACCGAGAGAACAGGTGTTTCATGAGGAAGTTCAAGATCGGGTACA 1846
Qy 1559 GCAACAACGGCTCGGACTTGGAAAGATGATCATGGATGACAGAAACGCAAGGCGAAGTCTT 1618
Dd 1847 GCAACAACGGCTCGGACTTGGAAAGATGATCATGGATGACAGAAACGCAAGGCGAAGTCTT 1906
Qy 1619 TTGAGGGCAACAACAACTATATACACCTGAGCTCGGACTTTTCCAGCTCTCTCCACGC 1678
Dd 1907 TTGAGGGCAACAACAACTATATACACCTGAGCTCGGACTTTTCCAGCTCTCTCCACGC 1966
Qy 1679 GATTCATCAGGATCTACCCCGAGAGACCACTCATGTCGGGACTGGGGCTCAGAAATGGAGC 1738
Dd 1967 GATTCATCAGGATCTACCCCGAGAGACCACTCATGTCGGGACTGGGGCTCAGAAATGGAGC 2026
Qy 1739 TGCTGGGCTGTGAAGTGGAAAGCCCTACAGCTGGACCCGACCACTCCCAACGGGAACTTGG 1798
Dd 2027 TGCTGGGCTGTGAAGTGGAAAGCCCTACAGCTGGACCCGACCACTCCCAACGGGAACTTGG 2086
Qy 1799 TGGATGAATGTATGACGACGACGCGCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGC 1858
Dd 2087 TGGATGAATGTATGACGACGACGCGCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGC 2146
Qy 1859 TCACAGGTGGCACCACTGTGCTGGCCACAGAAAGCCGCTCATAGACGACCACTATAC 1918
Dd 2147 TCACAGGTGGCACCACTGTGCTGGCCACAGAAAGCCGCTCATAGACGACCACTATAC 2206
Qy 1919 AATCAGAGTTTCCAAACATATGGTTTTTAACCTGTGAATTTTGGCTGGGGCTCTCAAAAGACT 1978
Dd 2207 AATCAGAGTTTCCAAACATATGGTTTTTAACCTGTGAATTTTGGCTGGGGCTCTCAAAAGACT 2266
Qy 1979 TCTGCCACTGGGAACATGACAAATCAGTGCAGCTCAAGTGGAGTGTGTGACCACCAAGA 2038
Dd 2267 TCTGCCACTGGGAACATGACAAATCAGTGCAGCTCAAGTGGAGTGTGTGACCACCAAGA 2326
Qy 2039 CGGGACCCATTCAGATCACACAGAGATGGCAACTTCATCTATTCCCAAGCTGACGAAA 2098
Dd 2327 CGGGACCCATTCAGATCACACAGAGATGGCAACTTCATCTATTCCCAAGCTGACGAAA 2386
Qy 2099 ATCAGAAGGGCAAACTGGCTCGCCTGGTAGCCCTGTGGTTTTATTTCCAGAACTCTGCC 2158
Dd 2387 ATCAGAAGGGCAAACTGGCTCGCCTGGTAGCCCTGTGGTTTTATTTCCAGAACTCTGCC 2446
Qy 2159 ACTGCATGACCTTCTGGTATACATGTCTGGTCCGCTCCAGCTCGGCACACTCAGGGTCAAA 2218
Dd 2447 ACTGCATGACCTTCTGGTATCACATGTCTGGTCCGCTCCAGCTCGGCACACTCAGGGTCAAA 2506
Qy 2219 TCCGCTACAGNAGCCAGGAGTACGATCAGCTGGTGGATGGCATTTGGACACCAAG 2278
Dd 2507 TCCGCTACAGNAGCCAGGAGTACGATCAGCTGGTGGATGGCATTTGGACACCAAG 2566
Qy 2279 GTGACCACTGGAAGGAAGGGCGTGTCTGTCCCAAGTCTCTGAAACTTTATCATAGGTGA 2338
Dd 2567 GTGACCACTGGAAGGAAGGGCGTGTCTGTCCCAAGTCTCTGAAACTTTATCATAGGTGA 2626
Qy 2339 TTTTCAGGGCGAAATCGGAAAGAAACCTTTGGTGGATTTGCTGGATGACATTAGTA 2398
Dd 2627 TTTTCAGGGCGAAATCGGAAAGAAACCTTTGGTGGATTTGCTGGATGACATTAGTA 2686
Qy 2399 TTAATAACCACTTTCACAAAGAAATTTGTGAAAACCAAGCAGACCTGGATAAAAAGAAC 2458
Dd 2687 TTAATAACCACTTTCACAAAGAAATTTGTGAAAACCAAGCAGACCTGGATAAAAAGAAC 2746

Db 901 |||||TGCTGCAGAGCGTCCCGCCTGTAACCTACCTGAGAATGGGTGGACTCCCGGAGAGGAT 960
Qy 961 TCCTACCGAGAGTGATACAGGTAGACTTGGCCCTTCTGCGCTTTGTACGCGCTCTCGGG 1020
Db 961 TCCTACCGAGAGTGATACAGGTAGACTTGGCCCTTCTGCGCTTTGTACGCGCTCTCGGG 1020
Qy 1021 ACACAGGGCGCCATTTCAAAAGAAAACCAAGAAAGAAATATTATGTCAAGACTTTACAAGATC 1080
Db 1021 ACACAGGGCGCCATTTCAAAAGAAAACCAAGAAAGAAATATTATGTCAAGACTTTACAAGATC 1080
Qy 1081 GACGTTAGTCTCAACGGGGAAGACTTGGATCACCATAAAGAAAGAAACAAACCTGTTCTC 1140
Db 1081 GACGTTAGTCTCAACGGGGAAGACTTGGATCACCATAAAGAAAGAAACAAACCTGTTCTC 1140
Qy 1141 TTTTACGAGAAACACCAACCCACAGATGTTGTGGTTCAGATATTCCTCCAAACCACTGATA 1200
Db 1141 TTTTACGAGAAACACCAACCCCTACAGATGTTGTGGTTCAGATATTCCTCCAAACCACTGATA 1200
Qy 1201 ACTCGATTGTGCCGAATCAAGCCTGCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1260
Db 1201 ACTCGATTGTGCCGAATCAAGCCTGCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1260
Qy 1261 GTATACGTTGCAAGATAACAGATTATCCTTGCTCTGGAATGTTGGGTATGGTCTCGGA 1320
Db 1261 GTATATGTTGCAAGATAACAGATTATCCTTGCTCTGGAATGTTGGGTATGGTCTCGGA 1320
Qy 1321 CTTATTCTGACTCCAGATCATCATCAACCAACCAAGGAGACAGAACTGGATGCTGNA 1380
Db 1321 CTTATTCTGACTCCAGATCATCATCAACCAACCAAGGAGACAGAACTGGATGCTGNA 1380
Qy 1381 AACATCCGCGCTGGTAACAGTGCCTCTGGCTGGGCATTTCCACCGGCACCTCATTTCCCTAC 1440
Db 1381 AACATCCGCGCTGGTAACAGTGCCTCTGGCTGGGCATTTCCACCGGCACCTCATTTCCCTAC 1440
Qy 1441 ATCAATGAGTGCTCCAAATAGACTTGGGGAGGAGAAAGATCGTGAGGGGCATCATATT 1500
Db 1441 ATCAATGAGTGCTCCAAATAGACTTGGGGAGGAGAAAGATCGTGAGGGGCATCATATT 1500
Qy 1501 CAGGTTGGGAACCGGAGAGAACAGGTGTTTCATGAGGAAGTTCAAGATCCGGGTACAGC 1560
Db 1501 CAGGTTGGGAACCGGAGAGAACAGGTGTTTCATGAGGAAGTTCAAGATCCGGGTACAGC 1560
Qy 1561 AACACGGCTCGGACTGGAAGATGATCATGGATGACAGCAAAACGGAAGCTTTT 1620
Db 1561 AACACGGCTCGGACTGGAAGATGATTCATGGATGACAGCAAAACGGAAGCTTTT 1620
Qy 1621 GAGGCAACAACAATATGATACACTGAGCTGCGGACTTTTCCAGCTCTCTCCACGGCA 1680
Db 1621 GAGGCAACAACAATATGATACACTGAGCTGCGGACTTTTCCAGCTCTCTCCACGGCA 1680
Qy 1681 TTTCATCAGGATCTACCCGAGAGAGCCACTCATGGGACTGGGCTCAGAATGAGCTG 1740
Db 1681 TTTCATCAGGATCTACCCGAGAGAGCCACTCATGGGACTGGGCTCAGAATGAGCTG 1740
Qy 1741 CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCGACCACTCCCAACGGGAACCTTGGTG 1800
Db 1741 CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCGACCACTCCCAACGGGAACCTTGGTG 1800
Qy 1801 GATGAATGTGATGACGACAGGCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
Db 1801 GATGAATGTGATGACGACAGGCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
Qy 1861 ACAGTGGCACCACCTGCTGCGCCACAGAAAAGCCCGGCTCATAGACAGCACCATACAA 1920
Db 1861 ACAGTGGCACCACCTGCTGCGCCACAGAAAAGCCCGGCTCATAGACAGCACCATACAA 1920
Qy 1921 TCAGAGTTTCCAAACATATGGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAAGACCTTC 1980
Db 1921 TCAGAGTTTCCAAACATATGGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAAGACCTTC 1980
Qy 1981 TCCCACTGGGAACATGACAATCACGTGCAAGTCAAGTGGAGTGTGTTGACCAGCAAGAGC 2040
Db 1981 |||||

Db 1981 TGCCACTGGGAACATGACATCACGTGACGCTCAAGTGGAGTGTGTTGACCAGCAAGACG 2040
Qy 2041 GGACCCATTTCAGGATCACA---CAGSAGATGSCAACTTCATCTATTCCCAAGCTCAGCAA 2097
Db 2041 GGACCCATTTCAGGATCACAACAGCAGGAGATGGCAACTTCACTATTTCCCAAGCTCAGCAA 2100
Qy 2098 AATCAGAAAGGCAAAAGTGGCTCGCCTGGTGGAGCCCTTGGTTTATTTCACAGAACTCTGCC 2157
Db 2101 AATCAGAAAGGCAAAAGTGGCTCGCCTGGTGGAGCCCTTGGTTTATTTCACAGAACTCTGCC 2160
Qy 2158 CACTGCATGACCTTCTGGTATCACAATGCTGGGTCCCAGGTGCGGCACACTCAGGTCAAA 2217
Db 2161 CACTGCATGACCTTCTGGTATCACAATGCTGGGTCCCAGGTGCGGCACACTCAGGTCAAA 2220
Qy 2218 CTGGCCTACCAAGACCAGAGGAGTACGATCAGCTGGTCTGGATGCCATTGGACACCAA 2277
Db 2221 CTGGCCTACCAAGACCAGAGAGTACGATCAGCTGGTCTGGATGCCCTTGGACACCAA 2280
Qy 2278 GGTGACCCTGGAAGGAGGCGTCTCTTGCTTCCACAAGTCTCTGAAACTTTTATCAGGTG 2337
Db 2281 GGTGACCCTGGAAGGAGGCGTCTCTTGCTTCCACAAGTCTCTGAAACTTTTATCAGGTG 2340
Qy 2338 ATTTTCGAGGGGAAATCGGAAAAGAAACCTTGGTGGGATTCCTGATGACATTAAT 2397
Db 2341 ATTTTCGAGGGGAAATCGGAAAAGAAACCTTGGTGGGATTCCTGATGACATTAAT 2400
Qy 2398 ATTAATAACCCACATTTCAACAAGAATTTGTGCAAAACAGCAGACCTGATATAAAGAAC 2457
Db 2401 ATTAATAACCCACATTTCAACAAGAATTTGTGCAAAACAGCAGACCTGATATAAAGAAC 2460
Qy 2458 CCAGAAATTAATAATGATGAAACAGGAGCAGCGCCAGGATGACGAAGTGAAGGAAAGT 2517
Db 2461 CCAGAAATTAATAATGATGAAACAGGAGCAGCGCCAGGATGACGAAGTGAAGGAAAGT 2520
Qy 2518 GACAAGACATCTCAGGAAGCCAGGCAATGTTGTAAGACCTTGAAGCCCATCCTCATC 2577
Db 2521 GACAAGACATCTCAGGAAGCCAGGCAATGTTGTAAGACCTTGAAGCCCATCCTCATC 2580
Qy 2578 ACCATCATAGCCATCAGCGCCCTGGGGTCCCTCGGGGCTGCTGTTGGGGTCTGCTG 2637
Db 2581 ACCATCATAGCCATCAGCGCCCTGGGGTCCCTCGGGGCTGCTGTTGGGGTCTGCTG 2640
Qy 2638 TACTGTGCTGTTGGCATAATGGGATGTCAGAAAAGAAACTTGTCTGCCCTGGAGAACTAT 2697
Db 2641 TACTGTGCTGTTGGCATAATGGGATGTCAGAAAAGAAACTTGTCTGCCCTGGAGAACTAT 2700
Qy 2698 AACTTTGAACCTTGTGGATGGTGAAGTTGAAAAGACAAACTGAATACACAGAGTACT 2757
Db 2701 AACTTTGAACCTTGTGGATGGTGAAGTTGAAAAGACAAACTGAATACACAGAGTACT 2760
Qy 2758 TATTCGAGGCAATGA 2772
Db 2761 TATTCGAGGCAATGA 2775

RESULT 8

AAF83311
ID AAF83311 standard; DNA; 2568 BP.

XX AAF83311;

AC AC
XX XX
DT 09-JUL-2001 (first entry)

XX Human NP-1 extracellular domain encoding DNA.

XX Receptor protein; vascular endothelial growth factor receptor-2;
KW VEGFR-2; neuropilin-1; NP-1; co-receptor; human; angiogenic; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
FH 1..2568
FT CDS /*tag= a

FT /product= "NP-1 extracellular domain"

WO200131346-A2.

03-MAY-2001

26-OCT-2000: 2000WO-IIS29579

28-QCT-1999. 99MS-0162367

(PROC \ PROCTER & GAMBLE CO

Rosenbaum JS, Whitaker GB, Limberg BJ:

WPT: 2001-308686/32.

P-PSDB: AAB62478.

Determining compounds which bind to a complex comprising vascular endothelial growth factor receptor-2 and Neuropilin-1 to provide superior pro- and anti-angiogenic agents -

endothelial growth factor receptor-2 and Neu-

superior pro- and anti-angiogenic agents

Example 7; page 55-61; 82pp; English.

The invention relates to determining whether a compound is capable of binding to a receptor protein complex comprising a vascular endothelial growth factor receptor-2 (VEGFR-2) receptor protein and a neuropilin-1 (NP-1) receptor protein. One method comprises introducing a sample comprising the compound to the receptor protein and allowing the compound to bind to the complex. Signaling through VEGFR-2 is enhanced in the presence of the NP-1 co-receptor. The methods of the invention can be used for identifying novel pro- and anti-angiogenic compounds. The present sequence represents the DNA encoding a human NP-1 extracellular domain.

The invention relates to determining whether a compound is capable of binding to a receptor protein complex comprising a vascular endothelial

binding to a receptor protein comprising a vascular endothelial growth factor receptor-2 (VEGFR-2) receptor protein and a neuropilin-1

(NP-1) receptor protein. One method comprises introducing a sample

comprising the compound to the receptor protein and allowing the

compound to bind to the complex. Signaling through VEGFR-2 is enhanced

in the presence of the NP-1 co-receptor. The methods of the invention can be used for identifying novel pro- and anti-apoptogenic compounds

The present sequence represents the DNA encoding a human NP-1

extracellular domain.

C

sequence 2308 BP; 120 A; 620 C; 644 G; 584 T; 0 other;

Query Match 91.5%; Score 2536.6; DB 22; Length 2568;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2555; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

```

every match
  best Local Similarity 99.58: score 2330.0
  best Local Similarity 99.58: Pred No. 0:

```

Category	Conservative	Mismatches	Indels	Gaps
Conservative	2555	0	9	3
Mismatches	2555	0	9	3
Indels	2555	0	9	3
Gaps	2555	0	9	3

QY 1 ATGGAGAGGGGCTGCCGCTCCTCTGGCCGCTGCTCGCCCTCGTCTCTCGCCCGGCCGC 60

db 1 ATGGAGAGGGGCTGCCGCTCCTCTGCGCCGTGCTGCCCTGCTCCTGCGCCGGCCGGC 60

QV 61 GCTTTTCGCAACGATGAATGTGGCGATACTATAAAATTGAAAGCCCGGGTACCTTACA 120

Db 61 GCTTTTCGCAACGATAAATGTGGCGATACTATAAAATTGAAAGCCCCGGGTACCTTACA 120

[illegible][illegible]

09
181

181

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

5
1
1

0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

Db 1561 AACACGGCTCGGACTGGAAGATGATTATGGATGACAGCAAAACGACGCGAAGTCTTTT 1620
QY 1621 GAGGCAACAACTATATACACTGAGCTGCGGACTTTTCCAGCTCTCTCCACGCA 1680
Db 1621 GAGGCAACAACTATATACACTGAGCTGCGGACTTTTCCAGCTCTCTCCACGCA 1680
QY 1681 TTTCATCAGATCTACCCGAGAGAGCCACTATGCGGAGTGGGCTCAGAAATGAGCTG 1740
Db 1681 TTTCATCAGATCTACCCGAGAGAGCCACTATGCGGAGTGGGCTCAGAAATGAGCTG 1740
QY 1741 CTGGCTGTGAAGTGAAGCCCTACAGCTGACCGACCACTCCCAAGGGAACTTGGT 1800
Db 1741 CTGGCTGTGAAGTGAAGCCCTACAGCTGACCGACCACTCCCAAGGGAACTTGGT 1800
QY 1801 GATGAATGTGATGACGACACGAGCCAACTGCCACAGTGGAAAGGTGATGACTTCCAGCTC 1860
Db 1801 GATGAATGTGATGACGACACGAGCCAACTGCCACAGTGGAAAGGTGATGACTTCCAGCTC 1860
QY 1861 ACAGTGGCACCACTGTGTGCTGCCACAGAAAGCCACGGTTCATAGACAGCACCATACAA 1920
Db 1861 ACAGTGGCACCACTGTGTGCTGCCACAGAAAGCCACGGTTCATAGACAGCACCATACAA 1920
QY 1921 TCAGATTTCCAAACATATGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAGACCTTC 1980
Db 1921 TCAGATTTCCAAACATATGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAGACCTTC 1980
QY 1981 TGCCACTGGGAACATGACATCAGCTGACGCTCAAGTGGAGTGTGTGACCAAGCAGAGC 2040
Db 1981 TGCCACTGGGAACATGACATCAGCTGACGCTCAAGTGGAGTGTGTGACCAAGCAGAGC 2040
QY 2041 GGACCCATTCAGGATCACA - - - CAGGAGATGGCAACTTCATCTATTCCCAAGCTGACGAA 2097
Db 2041 GGACCCATTCAGGATCACA CAGCAGAGATGGCAACTTCATCTATTCCCAAGCTGACGAA 2100
QY 2098 ATCAGAGGGCAAGTGGCTGCCCTGGTGAGCCCTGTGTTTATCCAGAACTCTGCC 2157
Db 2101 AATCAGAGGGCAAGTGGCTGCCCTGGTGAGCCCTGTGTTTATCCAGAACTCTGCC 2160
QY 2158 CACTGATGACCTTCTGGTATCACAATGCTGGTCCAGCTCGGCACACTCAGGTCAAA 2217
Db 2161 CACTGATGACCTTCTGGTATCACAATGCTGGTCCAGCTCGGCACACTCAGGTCAAA 2220
QY 2218 CTGGCTACCAAGACCCAGAGAGTACGATCAGCTGGTCTGGATGGCCATTTGGACACCAA 2277
Db 2221 CTGGCTACCAAGACCCAGAGAGTACGATCAGCTGGTCTGGATGGCCCTTGACACCAA 2280
QY 2278 GGTGACCACTGAAAGGAGGGCTGTCTGTCTCCACAAGTCTCTGAAACTTTATCAGGTG 2337
Db 2281 GGTGACCACTGAAAGGAGGGCTGTCTGTCTCCACAAGTCTCTGAAACTTTATCAGGTG 2340
QY 2338 ATTTTCGAGGGGAAATCGGAAAGGAAACCTTGGTGGATGCTGTGGATCACAATTAGT 2397
Db 2341 ATTTTCGAGGGGAAATCGGAAAGGAAACCTTGGTGGATGCTGTGGATCACAATTAGT 2400
QY 2398 ATTAATAACCACTTTACAAAGAAGATTGTGCAAAACCCAGCAGACCTGGATTAAGAAC 2457
Db 2401 ATTAATAACCACTTTACAAAGAAGATTGTGCAAAACCCAGCAGACCTGGATTAAGAAC 2460
QY 2458 CCAGAAATTAATTAATGATGAACACAGGAGCAGCCAGGATACAGAGTGAAGAGAGT 2517
Db 2461 CCAGAAATTAATTAATGATGAACACAGGAGCAGCCAGGATACAGAGTGAAGAGAGT 2520
QY 2518 GACAAAGAACCTCTCCAGGAAGCCAGGCAATGTTTGAAGACCTTAGA 2564
Db 2521 GACAAAGAACCTCTCCAGGAAGCCAGGCAATGTTTGAAGACCTTAGA 2567

RESULT 9
AAS06049
ID AAS06049 standard; DNA; 2808 BP.
XX
AC AAS06049;
XX

DT 12-SEP-2001 (first entry)
XX Angiotensin converting enzyme (ACEV) splice variant DNA #49.
XX Angiotensin converting enzyme splice variant; ACEV; Interleukin 6;
KW granulocyte colony stimulating factor receptor; glucagon; hypertrophy;
KW platelet-derived endothelial cell growth factor; cardiovascular disease;
KW cellular tumour antigen P53; cyclin-dependent kinase inhibitor 1C; ds;
KW vasoactive intestinal polypeptide receptor 2; arteriosclerosis; cancer;
KW myocardial infarction; coronary arterial thrombosis; renal disease;
KW diabetic nephropathy; muscular disease; immune disorder; sarcoidosis;
KW multiple sclerosis; immune complex nephritis; deep vein thrombosis;
KW nonaroidotic pulmonary granulomatous disease; endothelial abnormality;
KW vascular disorder; asbestosis.
XX Homo sapiens.
XX WO200136632-A2.
PN 25-MAY-2001.
XX 17-NOV-2000; 200WO-IL00766.
PF 17-NOV-1999; 99IL-0132978.
PR 10-DEC-1999; 99IL-0133455.
XX (COMP-) COMPUGEN LTD.
PA Levine Z, David A, Azar I, Khosravi R, Bernstein J;
PI WPI; 2001-336004/35.
DR P-PSDB; AAU02949.
XX Novel alternative splicing variants e.g. variant of angiotensin
PT converting enzyme (ACEV), useful in identifying candidate compounds
PT capable of binding to the variant and to detect anti-variant antibodies
XX Claim 1; Page 349-350; 519pp; English.
XX The sequence represents a DNA encoding an angiotensin converting enzyme
CC splice variant (ACEV) polypeptide. The polypeptides of the invention
CC include variants of granulocyte colony stimulating factor receptor,
CC glucagon, interleukin 6, platelet-derived endothelial cell growth factor,
CC cyclin-dependent kinase inhibitor 1C, cellular tumour antigen p53, and
CC vasoactive intestinal polypeptide receptor 2. The polypeptides and their
CC associated nucleic acids are useful for identification of variant
CC sequences and detection of candidate compounds capable of binding the
CC molecules. The sequences of the invention can be used in the treatment
CC and diagnosis of various disorders including cardiovascular diseases such
CC as arteriosclerosis, myocardial infarction and coronary arterial
CC thrombosis, renal diseases such as diabetic nephropathy, muscular
CC diseases such as hypertrophy, immune disorders such as immune complex
CC nephritis, multiple sclerosis, cancer, sarcoidosis, nonaroidotic
CC pulmonary granulomatous diseases such as asbestosis and vascular
CC pathologies involving an endothelial abnormality such as deep vein
CC thrombosis.
XX Sequence 2808 BP; 771 A; 667 C; 722 G; 646 T; 2 other;

Query Match 81.5%; Score 2258; DB 22; Length 2808;
Best Local Similarity 90.8%; Pred. No. 0;
Matches 2318; Conservative 2; Mismatches 2; Indels 250; Gaps 1;
QY 1 ATGGAGAGGGGCTGCGGCTCTCTGCGCGCTGCTGCGCGCTGCTGCGCGCGC 346
Db 287 ATGGAGAGGGGCTGCGGCTCTCTGCGCGCTGCTGCGCGCTGCTGCGCGCGC 346
QY 61 GCTTTTCCCAACGATGATGCGCATACTATAAAATTAAGAGCCCGGGTACCTTACA 120
Db 347 GCTTTTCCCAACGATGATGCGCATACTATAAAATTAAGAGCCCGGGTACCTTACA 406
QY 121 TCTCTGGTTATCTCTTATTCACCAAGTGAAGAAATGCGAATGGCTGATTCAGGCT 180

|||||
Db 407 TCCTCCTGGTATCTCTCATCTCTTATCATCCCAAGTGCAAAATGCGAATGGCTGATTTCAGGCT 466
|||||
QY 181 CCGGACCATACACAGAGAATATGATCAACTTCAACCCCTCAGTTCGATTTGGAGGACAGA 240
|||||
Db 467 CCGGACCATACACAGAGAATATGATCAACTTCAACCCCTCAGTTCGATTTGGAGGACAGA 526
|||||
QY 241 GACTGCAAGTATGACTACGAGTGAAGTCTTCGATGGAGAAATGAAATGACATTTTAGG 300
|||||
Db 527 GACTGCAAGTATGACTACGAGTGAAGTCTTCGATGGAGAAATGAAATGACATTTTAGG 586
|||||
QY 301 GGAAGTCTGTGGAGATAGCCCTCCCTCGTGTGTGTCCTCAGGGCCATTCCTTTT 360
|||||
Db 587 GGAAGTCTGTGGAGATAGCCCTCCCTCGTGTGTGTCCTCAGGGCCATTCCTTTT 646
|||||
QY 361 ATCAAAATTTCTCTGACTACGAAACACATGGTGCAGGATTTTCCATACGTTATGAAT 420
|||||
Db 647 ATCAAAATTTCTCTGACTACGAAACACATGGTGCAGGATTTTCCATACGTTATGAAT 706
|||||
QY 421 TTCAAGAGAGGTCTCTGAATGTTCCAGAACTACACAACTAGTGGAGTGATAAGTCC 480
|||||
Db 707 TTCAAGAGAGGTCTCTGAATGTTCCAGAACTACACAACTAGTGGAGTGATAAGTCC 766
|||||
QY 481 CCCGGATTCCTGAAATAATCCACAGACCTTGAATGCACTTATATGTCCTTTGGGCCA 540
|||||
Db 767 CCCGGATTCCTGAAATAATCCACAGACCTTGAATGCACTTATATGTCCTTTGGGCCA 826
|||||
QY 541 AAGATGTCAGAGATATCTCTGGAATTTGAAAGCTTTGACCTGGAGCTGACTCAAAATCCT 600
|||||
Db 827 AAGATGTCAGAGATATCTCTGGAATTTGAAAGCTTTGACCTGGAGCTGACTCAAAATCCT 886
|||||
QY 601 CCAGGGGGGATGTTCTCTCGCTACGACCGGCTAGAAATCTGGAGTGGATTCCTGATGT 660
|||||
Db 887 CCAGGGGGGATGTTCTCTCGCTACGACCGGCTAGAAATCTGGAGTGGATTCCTGATGT 946
|||||
QY 561 GGCCTTCATTTGGGGTTACTGTGACAGAAACACAGGTGCAATCCGATCCTCATCG 720
|||||
Db 947 GGCCTTCATTTGGGGTTACTGTGACAGAAACACAGGTGCAATCCGATCCTCATCG 1006
|||||
QY 721 GGCATCTCTCCATGTTTTCACCGCAGCGCGATAGCAAAAGATTTCTCAGCA 780
|||||
Db 1007 GGCATCTCTCCATGTTTTCACCGCAGCGCGATAGCAAAAGATTTCTCAGCA 1066
|||||
QY 781 AACTACAGTGTCTGCAGACAGTGTCTCAGAAAGATTTCAAAATGATGGAAGCTCTGGC 840
|||||
Db 1067 AACTACAGTGTCTGCAGACAGTGTCTCAGAAAGATTTCAAAATGATGGAAGCTCTGGC 1126
|||||
QY 841 ATGGAATCAGGAAATTCATTTGTACCCAGATCAGAGCTTCTCCAGTATAGCACCAC 900
|||||
Db 1127 ATGGAATCAGGAAATTCATTTGTACCCAGATCAGAGCTTCTTCCAGTATAGCACCAC 1186
|||||
QY 901 TGGTCTGCAGAGCGCTCCCGCTGAACTACCTGAGAAATGGGTGGACTCCCGGAGAGGAT 960
|||||
Db 1187 TGGTCTGCAGAGCGCTCCCGCTGAACTACCTGAGAAATGGGTGGACTCCCGGAGAGGAT 1246
|||||
QY 961 TCTTACCGAGTGGATACAGTGTGAGTGTGGGCTTCTGCGCTTGTCTCAGGCTGTGGG 1020
|||||
Db 1247 TCTTACCGAGTGGATACAGTGTGAGTGTGGGCTTCTGCGCTTGTCTCAGGCTGTGGG 1306
|||||
QY 1021 ACACAGGGGCCATTTCAAAAGAAACCAAGAGAAATATATGTCAAGACTTACAAGATC 1080
|||||
Db 1307 ACACAGGGGCCATTTCAAAAGAAACCAAGAGAAATATATGTCAAGACTTACAAGATC 1366
|||||
QY 1081 GAGTGTAGTCTCAACGGGGAAGTCTGATCACCATAAAAGAGAAACCAACCTGTCTC 1140
|||||
Db 1367 GAGTGTAGTCTCAACGGGGAAGTCTGATCACCATAAAAGAGAAACCAACCTGTCTC 1426
|||||
QY 1141 TTTACGGGAAACACCAACCCACAGATGTGTGGTGTGAGTATTTCCCAACCACTGATA 1200
|||||
Db 1427 TTTACGGGAAACACCAACCCACAGATGTGTGGTGTGAGTATTTCCCAACCACTGATA 1486
|||||
QY 1201 ACTCGATTTGTCGGAATCAAGCTTGAATGGGAACCTGGGCATATCTATGAGATTTGAA 1260
|||||

Db 1487 ACTCGATTTGTCGGAATCAAGCCTGCAACTTGGAAACTGGCATATCTATGAGATTTGAA 1546
QY 1261 GTATACGGTTGCAAGATAACAGATATTCTCTGCTGGAATGTGGGTATGGTCTCTGGA 1320
|||||
Db 1547 GTATACGGTTGCAAGATAACAGATATTCTCTGCTGGAATGTGGGTATGGTCTCTGGA 1606
|||||
QY 1321 CTTATTTCTGACTCCAGATACATCATCATCAACCAAGGAGACAGAACTGGATGCCCTGAA 1380
|||||
Db 1607 CTTATTTCTGACTCCAGATACATCATCAACCAAGGAGACAGAACTGGATGCCCTGAA 1666
|||||
QY 1381 AACATCGGCTGTGAACAGTCTCTGCTGGGACTTCCACCCGACCTCATTCCTAC 1440
|||||
Db 1667 AACATCGGCTGTGAACAGTCTCTGCTGGGACTTCCACCCGACCTCATTCCTAC 1726
|||||
QY 1441 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAAAGATCGTGAGGGGCATCATCAT 1500
|||||
Db 1727 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAAAGATCGTGAGGGGCATCATCAT 1786
|||||
QY 1501 CAGGGTGGGAAGCACCGAGAGAAAGGTTTTCATGAGGAAGTTCAAGATCGGGTACAGC 1560
|||||
Db 1787 CAGGGTGGGAAGCACCGAGAGAAAGGTTTTCATGAGGAAGTTCAAGATCGGGTACAGC 1846
|||||
QY 1561 AACAAAGGCTCGGACTGGAAGATGATCATGATGACAGCAAAAGCAAGGCGAAGTCTTTT 1620
|||||
Db 1847 AACAAAGGCTCGGACTGGAAGATGATCATGATGACAGCAAAAGCAAGGCGA----- 1898
QY 1621 GAGGGCAACAACAATATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCAGCGCA 1680
Db 1899 ----- 1898
QY 1681 TTCATCAGGATCTACCCCGAGAGAGCCACTCTATGGGGACTTGGGGTCTAGAATGGAGCTG 1740
Db 1899 ----- 1898
QY 1741 CTGGGCTGTGAAGTGAAGCCCCCTACAGCTGGACCAGCCACTCCCACGGGAACTTGGTG 1800
Db 1899 ----- 1898
QY 1801 GATGAATGTGATGACGACCAAGCGCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
Db 1899 ----- 1898
QY 1861 ACAGTGGGACCACTGTGTGGCCACAGAAAAGCCACGGTCTATAGACAGCACATACAA 1920
|||||
Db 1899 --AGTGGGACCACTGTGTGGCCACAGAAAAGCCACGGTCTATAGACAGCACATACAA 1956
QY 1921 TCAGAGTTCCTCAACATATGTTTAACTGTGAATTTGGCTGGGCTCTCACAGAGCTTC 1980
|||||
Db 1957 TCAGAGTTCCTCAACATATGTTTAACTGTGAATTTGGCTGGGCTCTCACAGAGCTTC 2016
QY 1981 TGCCACTGGGAACATGACAATACGTGCACTCAAGTGGAGTGTGTGACAGCAAGACG 2040
|||||
Db 2017 TGCCACTGGGAACATGACAATACGTGCACTCAAGTGGAGTGTGTGACAGCAAGACG 2076
QY 2041 GGACCATTCAGGATCACACAGGAGATGGCACTTCACTTATTTCCCAAGCTGACGAAAT 2100
|||||
Db 2077 GGACCATTCAGGATCACACAGGAGATGGCACTTCACTTATTTCCCAAGCTGACGAAAT 2136
QY 2101 CAGAAGGGCAAGTGGCTGGCTGGGTCCAGCTCGGCACTCGGCACTCGAGGTCGACCA 2160
|||||
Db 2137 CAGAAGGGCAAGTGGCTGGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 2196
QY 2161 TGCATGACCTTCTGTTATCACATGTCTGGTCCAGCTCGGCACTCGGCACTCGAGGTCGAACTG 2220
Db 2197 TGCATGACCTTCTGTTATCACATGTCTGGTCCAGCTCGGCACTCGAGGTCGAACTG 2256
QY 2221 CGCTACAGAACCCAGAGAGTACGATCAGTGGTCTGGATGGCCATTTGGACACCAAGGT 2280
Db 2257 CGCTACAGAACCCAGAGAGTACGATCAGTGGTCTGGATGGCCATTTGGACACCAAGGT 2316
QY 2281 GACCACTGGAAGGAGGGGCTGTCTTGTCTCCCAAGCTCTGAACTTTATCAGGTGAT 2340
Db 2317 GACCACTGGAAGGAGGGGCTGTCTTGTCTCCCAAGCTCTGAACTTTATCAGGTGAT 2376

QY 2341 TTCGAGGGCGAAATCGGAAAGGAAACCTTGCTGGGATTGCTGTGATGACATTAGTATT 2400
DB 2377 TTCGAGGGCGAAATCGGAAAGGAAACCTTGCTGGGATTGCTGTGATGACATTAGTATT 2436
QY 2401 AATAACCCACATTTTCAAGAAGATTGTGCAAAACAGCAGACCTGGATGATAAAGAACCCA 2460
DB 2437 AATAACCCACATTTTCAAGAAGATTGTGCAAAACAGCAGACCTGGATGATAAAGAACCCA 2496
QY 2461 GAAATTAATAATTGATGAACAGGGAGCAGCCAGGATACGAAGGTGAAGGAAAGGTGAC 2520
DB 2497 GAAATTAATAATTGATGAACAGGGAGCAGCCAGGATACGAAGGTGAAGGAAAGGTGAC 2556
QY 2521 AAGAATCTCTCAGGAAGCCAGGCAATGTGTTGAAGACCTTAGAACCCATCCTCATCACC 2580
DB 2557 AAGAATCTCTCAGGAAGCCAGGCAATGTGTTGAAGACCTTAGAACCCATCCTCATCACC 2616
QY 2581 ATCATAGCCATGAGCGCCCTGGGGTCTCTCGGGGGTCTGTGTGGGGTCTGTGTGATAC 2640
DB 2617 ATCATAGCCATGAGCGCCCTGGGGTCTCTCGGGGGTCTGTGTGGGGTCTGTGTGATAC 2676
QY 2641 TCTGCTGTGTCATAATGGGATGTGAGAAAGAAACTTGTCTGCCCTGGAGAACTATAC 2700
DB 2677 TGTGCTGTGTCATAATGGGATGTGAGAAAGAAACTTGTCTGCCCTGGAGAACTATAC 2736
QY 2701 TTTGAACCTTGTGGATGTGTAAGTTGAAAAAGACAACTGAATACAGAGTACTTAT 2760
DB 2737 TTTGAACCTTGTGGATGTGTAAGTTGAAAAAGACAACTGAATACAGAGTACTTAT 2796
QY 2761 TCGGAGGATGA 2772
DB 2797 TCGGAGGATGA 2808

RESULT 10

AA08418

ID AAX08418 standard; cDNA; 3471 BP.

XX AC AAX08418;

XX AC AAX08418;

XX DT 28-JUN-1999 (first entry)

XX DE Neuropilin gene.

XX DE Neuropilin gene.

KW Neuropilin; neuropilin-2; semaphorin; transmembrane protein; axon;

KW neuron; development; cell growth; immune response;

KW viral pathogenesis; treatment; disease; graft rejection;

KW viral disease; oncological disease; cancer; screening; probe; ss.

XX OS Rattus rattus.

XX FH Key

FT CDS Location/Qualifiers

FT 145..2913

FT /*tag= a

FT /product= "Neuropilin"

XX PN W09904263-Al.

XX PD 28-JAN-1999.

XX PF 17-JUL-1998; 98WO-US14632.

XX PR 17-JUL-1997; 97US-0052762.

XX PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.

XX PI Ginty DD, Kolodkin AL;

XX DR WPI; 1999-132446/11.

XX DR P-PSDB; AAW96309.

XX PT New isolated semaphorin receptor, neuropilin-2 - used to develop

PT products for the diagnosis and treatment of neurological,

PT immunological, oncological and viral diseases

XX Disclosure; Page 79-80; 88pp; English.

XX The neuropilins are type I transmembrane proteins and act as

CC semaphorin III (Sema III) receptors. The semaphorins have been

CC shown to function in repulsive axon guidance. Sema III is a

CC secreted protein that in vitro causes neuronal growth cone

CC collapse and chemorepulsion on neurites and is required in vivo for

CC correct sensory afferent innervation and other aspects of

CC development. Agents which inhibit or enhance the interaction of a

CC semaphorin and a neuropilin can be potent modulators of nerve cell

CC growth, immune responsiveness, and viral pathogenesis, and can be

CC used in the treatment and diagnosis of neurological disease,

CC neuro-regeneration, immune modulation including hypersensitivity and

CC graft-rejection, and diagnosis and treatment of viral and oncological

CC infection/diseases. The neuropilins, neuropilin-encoding nucleic

CC acids, and unique portions also are useful in screening chemical

CC libraries for regulators of semaphorin-mediated cell activity, and in

CC genetic mapping as probes for related genes, as diagnostic reagents

CC for genetic, neurological, immunological and oncological disease.

XX SQ Sequence 3471 BP; 931 A; 849 C; 912 G; 779 T; 0 other;

Query Match 79.7%; Score 2210.4; DB 20; Length 3471;

Best Local Similarity 87.6%; Pred. No. 0;

Matches 2428; Conservative 0; Mismatches 341; Indels 3; Gaps 1;

QY 1 ATGGAGAGGGGCTGCGGCTCTCTGCGCGCTGCTGCGCGCTGCTGCGCGCGCGCGCGCG 60

DB 145 ATGGAGAGGGGCTGCGGCTGCTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCGCGCGCG 204

QY 61 GCTTTTCGCAACGATGATGCGGATATATAAAATTTGAAAGCCCGGCTACCTTACA 120

DB 205 GCTTTTCGCGAGCATTAATGTCGCGGAGTATAAAATTTGAAAGCCCGGCTACCTTACA 264

QY 121 TCTCTGGTTATCTCTATTCATCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 180

DB 265 TCTCCCGGCTACCTTCTTACCATCCAAGTGAGAAATGTGAATGGCTAATCCAAGCT 324

QY 181 CCGGACCCATACCAGAAATTTATGATCACTTCAACCCCTCAGTTGAGTGAAGACAGA 240

DB 325 CCGGAGCCCTACAGAGAAATGATCACTTCAACCCCTCAGTTGAGTGAAGACAGA 384

QY 241 GACTCAAGTATGACTACGTGAAGTCTTCGATGAGAGAAATGAAATGGACATTTTAGG 300

DB 385 GACTCAAGTATGACTACGTGAAGTCTTCGATGAGAGAAATGAAATGGACATTTTAGG 444

QY 301 GGAAGTCTGTGGAAGATAGCCCTCTCTCTGTGTGTCTTCAGGGGCCATTTCTTTT 360

DB 445 GGAAGTCTGTGGAAGATAGCCCTCTCTCTGTGTGTCTTCAGGGGCCATTTCTTTT 504

QY 361 ATCAATTTGTCTGACTACCAACACATGGTGCAGGATTTTCCATAGTTATGAAAT 420

DB 505 ATCAATTTGTCTGACTACCAACACATGGTGCAGGATTTTCCATAGTTATGAAAT 564

QY 421 TTCAAGAGAGGTCTGAAATTTCCAGAACTACACAAACCTAGTGGAGTGATAAGTCC 480

DB 565 TTCAAGAGAGGTCTGAAATTTCCAGAACTACACAAACCTAGTGGAGTGATAAGTCC 624

QY 481 CCCGATTTCCCTGAAAAATATCCCAACGCTTGAATGCACCTTATATGCTTTTGGCCA 540

DB 625 CCTGGTTCCCTGAAAAATATCCCAACGCTTGAATGCACCTTATATGCTTTTGGCCA 684

QY 541 AAGATGTCAGAGATTTATCTGGAATTTGAAAGCTTTGACCTGGAGCTGACTCAATCT 600

DB 685 AAGATGTCAGAGATTTATCTGGAATTTGAAAGCTTTGACCTGGAGCTGACTCAATCT 744

QY 601 CCAGGGGGATGTTCTGTCGCTACGCGGCTAGAAATCTGGGATGGATTCCTGATGTT 660

DB 745 CCCGAGGAGTGTCTGTCGCTATGACCGGCTGGAGATCTGGGATGGATTCCTGATGTT 804

QY 661 GGCCTCATTGGGGCTTACTGTGGACAGAAAAACACAGGTGGAATCCGATCCTCATCG 720

Db 805 GGGCCCTCACATTTGGGCGTTTACTTGGGCGAAGAAATCTCTGGCGGATFCCGCTCCTCTTCA 864
QY 721 GGCATTCTCTCCATGTTTTTATACCGACAGCGCGATAGCAAAAGAGTTTCTCAGCA 780
Db 865 GGCATTCTATCCATGGTCTTCTACATGACAGCGCAATACAAAGGAGTTTCTCAGCC 924
QY 781 AACTACAGTGTCTGACAGACAGTGTCTCAGAAAGATTTCAAATGTATGAAAGCTCTGGGC 840
Db 925 AACTACAGCGTGTGACAGACAGCATCTCTGAAGATTTCAAAGTTATGAGGCTCTGGGC 984
QY 841 ATGGAATCAGGAGAAATTCATTCTGACCAGATCACAGCTTCTTCCCGAGTATAGCAACCA 900
Db 985 ATGGAATCTGGAGAGATFCCAATCTTGACACAGATCACTGCATCTTCCCAAGTATGTTACCA 1044
QY 901 TGTGCTCTCAGAGCGCTCCGCGCTGAATCACTCCCTGAGAAATGSGTGGAATCCCGGAGAGGAT 960
Db 1045 TGTGCTCTGAGCGCTCCCGCGCTGAATCACTCCCTGAAACGGGTGGACACAGGAGAGGAC 1104
QY 961 TCCTACGAGAGTGGATFACAGGTAGACTTGGGCGCTTCTGCGCTTTGTCAAGGCTGTGGGG 1020
Db 1105 TCCTACAGGAGTGGATFACAGGTGGACTTGGGCGCTCTCGGATTCGTTACTGCTGTGGGG 1164
QY 1021 ACACAGGCGCCATTTCCAAAGAAACCAAGAGAAATATTATGTCAAGACTTACAAGATC 1080
Db 1165 ACACAGGCTGCCATTTCCAAAGGAAACCAAGAGAAATATTATGTCAAGACTTACAGAGTA 1224
QY 1081 GAGTGTAGCTCCAAACGGGAGAGTGGATCACCATTAAAGAAAGAAACAACTGTCTC 1140
Db 1225 GACATCAGCTCCAAACGGAGAGGACTGGATCACCCTGAAGGAGGAAATAAAGCCATTATC 1284
QY 1141 TTTTCAGGAAACACCAACCCACAGATGTTGTGTGTGAGTATTCGCCAAACCACTGATA 1200
Db 1285 TTTTCAGGAAACACCAATCCACGGATGTGCTTTGGAGTTTTCGCCAAACCACTGATA 1344
QY 1201 ACTCGATTTCTCGAATCAAGCCTGCNACTTTGGGAACTGGCAATCTATGAGATTGAA 1260
Db 1345 ACTCGATTTCTCGAATCAAACTGCACTCTGGGAACTGGGAACTGGAATATCTATGAGATTGAA 1404
QY 1261 GTATAGGTTGCAAGATAACAGATTTATCCCTGCTCTGGAATGTTGGTATGTTGCTGGA 1320
Db 1405 GTTATGGCTGCAAGATAACAGATTTACCTTGTCTGGAATGTTGGCATGTTGCTGGA 1464
QY 1321 CTATTTCTGACTCCAGATFACATCATCAACCAAGGAGACAACTGGATGCCGTGAA 1380
Db 1465 CTATTTCTGACTCCAGATTTACAGCATCAACCAAGGAGACAGAACTGGATGCCAGAA 1524
QY 1381 AACATCGGCTGTAAACAGTCTGCTGCTGGGCGACTTCCACCGCACTCAFTCCCTAC 1440
Db 1525 AACATCGGCTGTGACCAAGTCAAGCGGCTGGGCGCTTGCACCCCTCACCCCAACCCATAC 1584
QY 1441 ATCAATGAGTGGCTCCAAATFAGACCTGGGGAGGAGAAATCTGTAGGGGCATCATCAT 1500
Db 1585 ATCAATGATGGCTCCAAAGTGGACCTGGGAGATGAGAGATAGTAGAGGTGTCAATCAT 1644
QY 1501 CAGGTTGGGAAGCACCGAGAGAAACAAGGTGTTTCATGAGGAAGTTCAAGATCGGGGTACAGC 1560
Db 1645 CAAGTTGGGAAGCACCGAGAGAAACAAGTGTTCATGAGGAAGTTCAAGATCGGCTACAGT 1704
QY 1561 AACAAAGGCTCGACTGGAGAGATCATGGATGACAGCAAGCAAGCGGAAGTCTTTT 1620
Db 1705 AACAAATGGTCTGACTGGAAATGATCATGGATGACAGCAAGCGCAAGGCTAAGTCTTTT 1764
QY 1621 GAGGCGCAACAACATATGATACACCTGAGCTGCGGAGTTCCTACGCTCTCTCCACGCGA 1680
Db 1765 GAAGGCAACAACAACATATGACACACCTGAGCTCCGGGCGCTTTACACCTCTCTCCACAAGA 1824
QY 1681 TTCATCAGGATCTACCCCGAGAGAGCCACTCATGCGGAGCTGGGGCTCAGAAATGGAGCTG 1740
Db 1825 TTCATCAGGATCTACCCCGAGAGAGCCACACATAGTGGGCTCGGACTGAGGATGGAGCTA 1884
QY 1741 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACCGACCACTCCCAACCGGAACCTTGGTG 1800
|||||

Db 1885 CTGGGCTGTGAAGTAGAAGTGCCTTACAGCTGGACCCACAGCACCCCAATGGGAACCCCGTG 1944
QY 1801 GATGAATGTGATACGACACAGGGCAACTGCCACAGTGGAAACAGAGTGATGACTTCCAGCTC 1860
Db 1945 GACGAGTGTGACGATGACAGGCGCAACTGCCACAGTGGCAGAGTGATGACTTCCAGCTC 2004
QY 1861 ACAGGTGGCACCACTGTGTGGCCACAGAAAGCCACGGTCTATAGACAGACCATCAAA 1920
Db 2005 ACAGGAGGCACTGTCTTGGCCACAGAGAGCAACCATTTATAGACAGCACCATCAA 2064
QY 1921 TCAGAGTTTCCACATATGTTTTAACTGTGAATTTGGCTGGGGCTCTCACAAGACCTTC 1980
Db 2065 TCAGAGTTTCCGACATACGTTTTAACTGCGAGTTTGGCTGGGGCTCTCACAAGACATTC 2124
QY 1981 TGGCACTGGCAACATGACAAATCAGTGCAGCTCAAGTGGAGTGTGTTGACCAGCAAGACG 2040
Db 2125 TGGCACTGGGAACATGACAGCCACGCGAGCTCAGTGGAGGGTGTGACACGCAAGACG 2184
QY 2041 GGACCCATTTCAGGATCACACAGGAGATGGCAACTTTCATCTATTCCCAGCTGACGAAAT 2100
Db 2185 GGCGCCATTTCAGGACACACAGGAGATGGCAACTTTCATCTATTCCCAGCTGATGAAAT 2244
QY 2101 CAGAAAGGCAAGTGGCTCGCTGGTGAGCCCTGTGTTTTATTTCCCAGAACTCTGCCCCAC 2160
Db 2245 CAGAAAGGCAAGTAGCCCGCTGGTGAGCCCTGTGTTCTATTCCCAGAGTTCTTGCCCCAC 2304
QY 2161 TGCATGACCTTCTGGTATCACATGCTCGGTCGCCACGTCGCGCACACTCAGGTCACAACTG 2220
Db 2305 TGCATGACCTTCTGGTATCACATGCTCGGCTCTCATGTGGGTACACTGAGGGTCAAACTG 2364
QY 2221 CGCTACAGAAAGCCAGAGGATACGATCAGCTGCTGTGGATGGCCATTTGACACACCAGGT 2280
Db 2365 CACTACAGAAAGCCAGAGGAATATGATCAACTGCTGTGGATGGTTCGGCCACCAGGA 2424
QY 2281 GACCACTGGAAGAAAGGCGCTGCTTGTCTCCCAAGTCTCTGAACTTTATCAGGTGATTT 2340
Db 2425 GACCACTGGAAGAAAGGCGCTGCTTGTCTGCACAAATCTCTGAACTGTATCAGGTATT 2484
QY 2341 TTGAGGGCGCAATTCGAAAGGAAACCTTTGGTGGGATGCTGTGGATGACATTAGTATT 2400
Db 2485 TTTTGAAGTGAATTCGAAAGAAAGAACCTCGTGGGATGCTGTGGATGATATCAGTATT 2544
QY 2401 AATAACCAATTTTCAAGAAAGATTGTGCAAAACACAGCACCTGGATGGAATAAAAGAACCA 2460
Db 2545 AACAAACCAATTTCTCAGGAGGACTGTGCAAAACCAACAGACACCTAGATAAAAGAACACA 2604
QY 2461 GAAATTAATTTATGAAACAGGAGGACGCCAGGATACGAAAGGTGAAGGAGAGGTGAC 2520
Db 2605 GAAATTAATATGAAACAGGAGGACGCCAGGATATGAA---GAAGGAAAGGCGAC 2661
QY 2521 AAGAACATCTCCAGGAGGACGAGCAATGTGTTGAAGACCTTAGAACCCATCTCATCACCC 2580
Db 2662 AAGAACATCTCCAGGAGGACGAGCAATGTGTTAAGACCTTGACCCCATCTGTATCACCC 2721
QY 2581 ATCATAGCCATGAGCGCCCTGGGGGCTCTCTCGGGGCTGTCTGTGGGGTGTGCTGTGTAC 2640
Db 2722 ATCATAGCCATGAGTGCCTGGGGTGTCTCTGGGTGCACTGTGTGAGTGTGTGTGTAC 2781
QY 2641 TGTGCTCTGTGGCATTAATGGGATGTCAGAAAGAAACTTGTCTGCCCTGGGAGAACTATAAC 2700
Db 2782 TGTGCTCTGTGGCACAATGGGATGTCGGAAGGAACCTTCTGCCCTGGGAGAACTATAAC 2841
QY 2701 TTTTGAATCTGTGGATGTTGAAGTTTGAAGGAGCAAACTGAAATACACAGAGTACTTAT 2760
Db 2842 TTTTGAATCTGTGGATGTTGAAGTTTGAAGGAGTAATAACTGAACCCACAGAGTAATTAC 2901
QY 2761 TCGAGGCGATGA 2772
Db 2902 TCAGAGGCGTGA 2913
RESULT 11
AAX08995

1608 GTTATGGCTGCAAGATAACAGATTATCTGCTGTGGAATGTTGGGCATGGTGTCTGGA 1667
1321 CTTATTCTGACCTCCAGATACATCATCAACCAAGGAGACAGAACTGGATGCCCTGAA 1380
1668 CTTATTTCAGACTCCAGATTACAGATTCATCAAGCCAGAGAAATGGATGCCAGAA 1727
1381 AACATCCGCTGTGAACCACTGCTCTGGCTGGGCACTTCCACCCGACCTCATTTCTTAC 1440
1728 AACATCCGCTGTGTGACCACTGTACCGGTGGGCACTGCCACCTCACCCACCCATAC 1787
1441 ATCAATAGTGGCTCCAAATAGACTGTGGGAGGAGAGAACTGTGAGGGGCATCATATT 1500
1788 ACCAATGAATGGCTCCAAAGTGGACCTGGGAGATGAGAAGATAGTAAGAGGTGTCACTATT 1847
1501 CAGGTTGGGAAGCACCGAGAGAAAGGTGTTTCATCAGAGAAAGTTCAAGATCGGGTACAGC 1560
1848 CAGGTTGGGAGACCCGAGAAACAAGGTGTTTCATGAGGAAGTTCAAGATCGGCTATAGT 1907
1561 AACACGGCTCGGACTGGAGATGATCATGGATGACAGCAAGCAAGCGCAAGTCTTTT 1620
1908 AACATGGCTCTGACTGGAAACTATCATGGATGACAGCAAGCGCAAGCTAAGTCGTTT 1967
1621 GAGGCAACAACAATATGATACACTGAGCTCGGCACTTTTCCAGCTCTCTCCACGGCA 1680
1968 GAAGGCAACAACAATATGATACACTGAGCTCTCGGCACTTTTTCACCTCTCTCCACAAG 2027
1681 TTCATCAGGATCTACCCGAGAGAGGACCTCATGCGGGGCTGAGGATCGGAGCTG 1740
2028 TTCATCAGGATCTACCCGAGAGAGGACCTCATGCGGGGCTGAGGATCGGAGCTG 2087
1741 CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCCAGCTCCCAAGCGGAATTTGGTG 1800
2088 CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCCAGCTCCCAAGCGGAATTTGGTG 2147
1801 GATGAATGTGATGACACAGGACCACTGCGCAGTGGACAGGTGATGACTTCCAGCTC 1860
2148 CATGAGTGTGACGACAGCAGGACCACTGCGCAGTGGACAGGTGATGACTTCCAGCTC 2207
1861 ACAGTGGCAGCACTGTCTGGCCACAGAAAGCCACCGTTCATAGACAGCAAGTACAA 1920
2208 ACAGGAGGACCACTGTCTGGCCACAGAGAGCCAACTATATAGACAGCAAGTACAA 2267
1921 TCAGATTTCCAAACATATGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAGACCTTC 1980
2268 TCAGATTTCCCAACATATGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAGACATTC 2327
1981 TGCCACTGGGACATGACAAATCACGTGACGCTCAAGTGGAGTGTGTTGACCAAGACG 2040
2328 TGCCACTGGGACATGACAAATCACGTGACGCTCAAGTGGAGTGTGTTGACCAAGACG 2387
2041 GGACCCATTCAGGATCACACAGAGATGGCACTTCATCTATTCCCAAGCTGACGAAAT 2100
2388 GGGCCGATTTCAGGACATACAGAGATGGCACTTCATCTATTCCCAAGCTGATGAAT 2447
2101 CAGAAGGGCAAGTGGCTCGCTGGTGGAGCCCTGTGGTTTATTCACAGAACTTGCCCCAC 2160
2448 CAGAAGGGCAAGTGGCTCGCTGGTGGAGCCCTGTGGTTTATTCACAGAACTTGCCCCAC 2507
2161 TGCCACTGGGACATGACAAATCACGTGACGCTCAAGTGGAGTGTGTTGACCAAGACG 2220
2508 TGATGACCTTCTGGTATCACATGCTCGGTGCCAGTCCAGTCCGACACTCAGGTCRAACTG 2280
2221 CGCTACAGAGCCAGGAGTACGATCAGTGGTGTGGATGGCCATTTGGACACCAAGGT 2280
2568 CGCTACAGAGCCAGGAGTATGATCACTGGTGTGGATGGTGGTGGGACCAAGGA 2627
2281 GACCACTGGAAGAGGGCGTGTCTGCTCCCAAGTCTCTGAACTTTATCAGGTGATT 2340
2628 GACCACTGGAAGAGGGCGTGTCTGCTCCCAAGTCTCTGAACTTTATCAGGTGATT 2687
2341 TTCGAGGGCGGAATCGAAGAGGAACCTTGGTGGGATGCTGGATCACATTAGTATT 2400

2688 TTTGAAGGTGAATCGAAGAAAGAAACCTTGGTGGAAATGCTGTGGATGATATCAGTATT 2747
2401 AATAAACACATTTCAAGAAGATTGTGCAAAAACACAGACCTGGATAAAAAAGAACCCA 2460
2748 AACACCATATTTCTCAGGAGACTGTGCAAAACACAGACCTAGATAAAAAGAACACA 2807
2461 GAAATTTAAATTTGATGAAACAGGAGCAGCCAGGATACGAAGGTGAAGGAGAAGGTGAC 2520
2808 GAAATTTAAATTTGATGAAACAGGAGCAGCTCCAGGATATGAAGGAGAAGGAGGTGAC 2867
2521 AGAACATCTCCAGGAAGCCAGCAATGTGTTGAAGACCTTAGAACCTCTCTCATCACC 2580
2868 AGAACATCTCCAGGAAGCCAGCAATGTGTTGAAGACCTTAGAACCTCTCTCATCACC 2927
2581 ATCATAGCATGAGCCCTCGGGGCTCTCTGGGGCTGTCTGTGGGTGCTGTGCTGTAC 2640
2928 ATCATAGCATGAGCCCTCGGGGCTCTCTGGGGCTGTCTGTGGGTGCTGTGCTGTAC 2987
2641 TGTGCTGTGGCATTAATGGGATGTCAGAAAGAAACTTGTCTGCCCTGGAGAACTATAAC 2700
2988 TGTGCTGTGGCATTAATGGGATGTCAGAAAGAAACTTGTCTGCCCTGGAGAACTATAAC 3047
2701 TTTGAATTTGTGGATGGTGTGAAGTTGAAAGAGAAACTGAATACACAGACTATTAT 2760
3048 TTTGAATTTGTGGATGGTGTGAAGTTGAAAGAGAAACTGAATACACAGACTATTAT 3107
2761 TCGGAGGCGATGA 2772
3108 TCAGAGGCGTGA 3119
RESULT 12
AAZ31430
ID AAZ31430 standard; cDNA; 3652 BP.
XX AAZ31430;
AC AAZ31430;
DT 07-FEB-2000 (first entry)
XX Mouse neuropilin cDNA sequence.
DE Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; mouse; ss.
XX Mus sp.
XX WO955855-A2.
PN 04-NOV-1999.
PD 23-APR-1999; 99WO-CA00324.
PF 23-APR-1999; 98US-0082791.
PR (GENE-) GENESENSE TECHNOLOGIES INC.
PA Wright JA, Young AH, Lee YS;
XX WPI; 2000-023357/02.
DR Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -
XX Disclosure; Fig 7; 57pp; English.
PS The invention provides sequences AAZ31431-460 which are antisense
CC oligonucleotides that inhibit human neuropilin expression. The antisense
CC oligonucleotides can be used to inhibit the growth or metastasis of a
CC mammalian tumor and inhibit neovascularisation. The oligonucleotides may
CC be used to treat various forms of cancers or tumors, such as sarcomas,
CC melanomas, adenomas, carcinomas of solid tissue, hypoxic tumors,
CC squamous cell carcinomas of the mouth, throat, larynx and lung,
CC genitourinary cancers such as cervical and bladder cancer, hematopoietic

CC cancers, colon cancer, breast cancer, pancreatic cancer, renal cancer,
CC brain cancer, skin cancer, liver cancer, head and neck cancers, and
CC nervous system cancers, as well as benign lesions such as papillomas. The
CC methods may be used to treat neovascularisation disorders such as
CC diabetic retinopathy, and retinopathy of prematurity and age related
CC macular degeneration. The present sequence represents the mouse
CC neuroplln cDNA.
xx
SQ

Sequence 3652 BP; 965 A; 874 C; 941 G; 872 T; 0 other;

Query Match 79.5%; Score 2202.4; DB 21; Length 3652;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 2416; Conservative 0; Mismatches 356; Indels 0; Gaps 0;

Qy 1 ATGGAGAGGGGCTCCCTCTCTGCGCGTCTGCGCCCTGCTCGCCCTGCGCCGCGCGCGG 60
Db 348 ATGGAGAGGGGCTCCCTCTCTGCGCGCTGCGCCCTGCGCCCTGCGCCCTGCGCGGCG 407
Qy 61 GCTTTTCGAACGATGAATGTTGGGATATATAAAATTTGAAAGCCCGGGTACCTTACA 120
Db 408 GCTTTTCGACGACAAATGTGGGGGACCATAAAAATCGAAACCCAGGGTACCTACA 467
Qy 121 TCTCTGCTTATCTCTATCTATACCCCAAGTGAAATGCGATGCGCTGATTCAGGCT 180
Db 468 TCTCCGGTTTACCTCTATCTTACCATCCAAAGTGAGAACTGTGAATGGCTAATCCAA 527
Qy 181 CCGGACCCATACCAGAGAAATATGATCAACTCAACCTCACTTCGATTTCGGAGACAGA 240
Db 528 CCGNACCCTACCAGAGAAATCAATCAACTCAACCCACATTTTCGATTTCGGAGACAGA 587
Qy 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGAGAGAAATGAAATGGACATTTTAG 300
Db 588 GACTGCAAGTATGACTACGTGGAAGTATTCGATGGGAGAAATGAAGGCGCGCTGTGG 647
Qy 301 GGAAAGTTCTGTGNAAGATAGCCCCCTCTCTGTGTGTCTTCAGGGCCATTTCTTTT 360
Db 648 GGGAAAGTTCTGTGGAAGATGCACTTCTCTGTGTGTCTTCAGGGCCCTTCTCTTC 707
Qy 361 ATCAAAATTTGTCTGTACTACGAAACACATGGTGAGAGATTTTCATACGTTATGAAAT 420
Db 708 ATCAAAATTTGTCTGTACTATGACACATGCGGACAGGGTTTTCATCCCGTATGAATC 767
Qy 421 TTCAGAGAGGTCTGAAATGTTCCAGAACTACACACACCTAGTGGAGTGATAAGTCC 480
Db 768 TTCAGAGAGGGCCGGAATGTTCTCAGAACTATACAGCACCTACTTGGAGTGATAAGTCC 827
Qy 481 CCCGATTCCTGAAAAATATCCCAACAGCCTTGAATGCATATATTTGCTTTGGCGCA 540
Db 828 CTTGGGTTCCCTGAAAAATACCCCACTGCTTGGAGTGCACCTACATCATCTTTGCACCA 887
Qy 541 AAGATGTCAGAGATTATCTCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAACT 600
Db 888 AAGATGTCAGAGAAATCTCTGGAGTTTGAAGTTTGGACCTGGAGCAAGACTCGAATCCT 947
Qy 601 CCAGGGGGAATGTTCTGTCTAGCAGCGGCTAGAAATCTGGGATGGAATCCCTGATGT 660
Db 948 CCCGAGGAATGTTCTGTCTGATGACCGGCTGGAGATCTGGGATGGAATCCCTGAACT 1007
Qy 661 GGCCCTCACATTTGGCGTTACTGTGGACAGAAACACACAGGTGCAATCCGATTCCTATCG 720
Db 1008 GGCCCTCACATTTGGCGGTTATTTGGGCGAGAAATCTCTGGCGGATCCCGCTCTCTTCA 1067
Qy 721 GGCATTCTCCATGGTTTTTACACGACAGCGGATAGCAAGAGGTTTCTCAGCA 780
Db 1068 GCGGTTCTATCCATGGTTTATACACTGACAGCGCAATAGCAAGAGGTTTCTCAGCC 1127
Qy 781 AACTACAGTGTCTTCGACAGCAGTGTCTCAGAAGATTTCAAAATGATGGAAGCTTGGCG 840
Db 1128 AACTACAGTGTCTTCGACAGCAGCATCTCTGAAGATTTTAAGTGTATGGAGCTCTGGCG 1187
Qy 841 ATGGAATCAGGAGAAATTCATTTCTGACAGATTCACAGCTTCTTCCCATATAGCACCAAC 900
Db 1188 ATGGAATCTGGAGATCCATTTCTGATCAGATCACTGCATCTTCACAGTATGGTACCAAC 1247

Qy 901 TGGTCTGCAGAGCGCTCCGCGCTGAACCTACCTTGAGATGGTGGACTCCCGGAGAGAT 960
Db 1248 TGGTCTGTAGAGCGCTCCGCGCTGAACCTACCTGAAATGGTGGACTCCAGGAGAGAC 1307
Qy 961 TCCACCGAGAGTGGATACAGGTAGACTTGGGCGCTTCTGCGCTTTGTCACGGCTGTCTGG 1020
Db 1308 TCCACAGAGAGTGGATCCAGGTGGACTTGGGCGCTTCTGCGATTCGTTACTGCTTAGGG 1367
Qy 1021 ACAGGGCGCCATTTCAAAAGAAACCAAGAAATATATATGCAAGACTTACAAGATC 1080
Db 1368 ACAGGGTGGCCATTTCCAAGGAACCAAGAAATATATATGCAAGACTTACAGATG 1427
Qy 1081 GACGTTAGCTCCAAAGGGAAGACTGGATACCATATAAAGAAAGAAACAAACCTGTTCTC 1140
Db 1428 GACATCAGCTCCAAAGGAGAGGACTGGATCTCCCTGAAAGAGGAAATAAAGCCATTATC 1487
Qy 1141 TTTCAGGGAACACACACCCACAGATGTTGTTGGTTGCAGTATTTCCCCAAACCACTGATA 1200
Db 1488 TTTCAGGGAACACACACCCACAGATGTTGTTCTTAGGAGTTTCTCCAAACCACTGATA 1547
Qy 1201 ACTCGATTTGTCGGAATCAAGCCTGCAACTTGGGAACTGCATATCTATGAGATTTGAA 1260
Db 1548 ACTCGATTTGTCGGAATCAAACTGTATCTCTGGGAACTGGTATATCTATGAGATTTGAA 1607
Qy 1261 GTATACGGTTTCAAGATAACAGATTATCTTGTCTCTGGAATGTTGGGTATGTGTCTCGA 1320
Db 1608 GTTATGCTGCAAGATAACAGATTATCTTGTCTCTGGAATGTTGGGATGTTGTCTCGA 1667
Qy 1321 CTTATTTCTGACTCCAGATCACATCATCAACCAAGAGAGACAGAACTGGATGCTCGAA 1380
Db 1668 CTTATTTCTGAGACTCCAGATTACAGATCAATCAAGCCGACAGGAATTTGGATGCGAGAA 1727
Qy 1381 AACATCCGCTGGTAACAGTCTGCTGCTGGCTGGGCACTTCCACCCGCACTCATTTCTCTAC 1440
Db 1728 AACATCCGCTGGTGACCACTGCTGCGGCTGGGCACTTCCACCCGCACTCATTTCTCTAC 1787
Qy 1441 ATCAATGATGGCTCCAAATAGACTGGGGAGGAGAAAGATCGTGAGGGGATCATCATTT 1500
Db 1788 ACCAATGATGGCTCCAAAGTGGACCTGGAGATGAGAGATAGTAAGAGGTGTCTCATTT 1847
Qy 1501 CAGGTTGGGAAGCAGGAGAAAGGTGTTTCATGAGGAAGTTCAAGATCGGTTACAGC 1560
Db 1848 CAGGTTGGGAAGCAGGAGAAAGGTGTTTCATGAGGAAGTTCAAGATGCGCTATAGT 1907
Qy 1561 AACAAAGCTCGGACTGGAAGATGATCATGATGACACAAACCAAGCCGCAAGTCTTTT 1620
Db 1908 AACAAAGCTCTGACTGGAAATATCATGATGACACAAAGCCGCAAGGCTTAAGTCGTT 1967
Qy 1621 GAGGCAACAACTATGATACACTGAGCTGGGCACTTTTCCAGCTCTCTCCACGCGA 1680
Db 1968 GAAGCAACAACTATGACACACTGAGCTTGGAGCTTCCAGCTCTCTCCACAGG 2027
Qy 1681 TTCATCAGGATCTACCCCGAGAGAGCCACTCATGGCGACTGGGGCTCAGAAATGGAGCTG 1740
Db 2028 TTCATCAGGATCTACCCCTGAGAGAGCCACACAGCTGGGCTTGGGCTGAGGATGGAGCTA 2087
Qy 1741 CTGGGCTGTAAGTGGAAAGCCCCCTACAGCTGGACCGACCTCCCAAGCGGAATTTGGTG 1800
Db 2088 CTGGGCTGTAAGTGGAAAGCCCCCTACAGCTGGACCGACCTCCCAAGCGGAATTTGGTG 2147
Qy 1801 GATGAATCTGATGACGACCGCCAACTGCCACAGTGGCAAGAGGTGATGACTTCCAGCTC 1860
Db 2148 CATGAGTGTGACGACGACCGCCAACTGCCACAGTGGCAGAGGTGATGACTTCCAGCTC 2207
Qy 1861 ACAGTGGCACCACTGTGCTGGCCACAGAAAGCCCGGTATAGACAGACCACTACAA 1920
Db 2208 ACAGGAGCACCACTGTCTCTGGCCACAGAGAAAGCAACATTATAGACAGCACTCA 2267
Qy 1921 TCAGAGTTTCAACATATGTTTTTAAGTGTGAATTTGGCTGGGCTCTCACAAGACCTTC 1980
Db 2268 TCAGAGTTTCCGAGATACGGTTTTAACTGCGAGTTTGGCTGGGCTCTCTCACAAGACATTC 2327

|||||
Db 2988 TGGGCTGTTGGCACAATGGGATGTGAGAAAGAACTATCTGCCCTGGAGAACTATAAC 3047
QY 2701 TTTGAACCTTGTGGATGGTGTGAAGTTGAAAAAGACAAACTGAATAACACAGAGTACTTAT 2760
Db 3048 TTTGAACCTTGTGGATGGTGTGAAGTTGAAAAAGATAAAGTAACCTGAACCCACAGAGTAATTAC 3107
QY 2761 TCGGAGGCATGA 2772
Db 3108 TCAGAGCGGTGA 3119
|| ||||| |||

RESULT 14

AAX08994
ID AAX08994 standard; cDNA; 2766 BP.
XX
AC AAX08994;
XX
DT 12-MAY-1999 (first entry)
DE
XX Rat semaphorin receptor 1 (SR1) coding sequence.

XX Semaphorin receptor; SR; regulation; cell function;
KW cell morphology; nerve cell; neuron; axon; diagnosis; therapy;
KW probe; primer; detection; mutant allele; neuropilin; ds.
XX

OS Rattus rattus.
XX

XX Key Location/Qualifiers
FH 1..2766
FT CDS /*tag= a
FT /product= SR1_polypeptide
XX

XX W0902556-A1.
PN

XX 21-JAN-1999.
PD

XX 08-JUL-1998; 98WO-US14290.
XX

XX 24-SEP-1997; 97US-0936135.
PR

XX 08-JUL-1997; 97US-0889458.
PR

XX (REG) UNIV CALIFORNIA.
XX

XX Chen H, He Z, Tessier-lavigne M;
PI

XX WPI; 1999-120780/10.
DR

XX New semaphorin receptor (SR) class 1 and class 2 proteins - useful
PT for, e.g. regulating nerve cell function and morphology

XX Claim 3; Page 48-49; 105pp; English.
PS

XX Semaphorin receptor (SR) proteins are involved in nerve cell
CC guidance. The SR proteins can regulate cell function and morphology,
CC especially nerve cells. Inhibitors of SR proteins promote nerve cell
CC growth. Binding agents specific to the SR polypeptides are useful
CC for diagnosis and therapy, especially where disease or disease
CC prognosis is associated with improper or undesirable axon outgrowth
CC or orientation. The SR nucleic acid sequences are used to provide
CC translatable transcripts, hybridisation probes, PCR primers,
CC diagnostic nucleic acids, etc. They are used in detecting the
CC presence of SR genes, and to identify wild type and mutant alleles.
CC Mutant alleles are used to generate allele-specific oligonucleotide
CC probes for clinical diagnosis. The methods are used to identify
CC agents which modulate the activity of SR polypeptides. This
CC sequence encodes the rat SR1 polypeptide.

XX Sequence 2766 BP; 766 A; 680 C; 716 G; 604 T; 0 other;
SQ

Query Match 79.2%; Score 2196; DB 20; Length 2766;
Best Local Similarity 87.5%; Pred. No. 0;

Matches 2426; Conservative 0; Mismatches 340; Indels 6; Gaps 2;

QY 1 ATGAGAGGGGGCTGCCGTCTCTCGCGCGTGTCTCGCCCTCGTCTCGCCCGCGCGCGC 60
Db 1 ATGAGAGGGGGCTGCCGTCTCTCGCGCGTGTCTCGCCCTCGTCTCGCCCGCGCGCGC 57
QY 61 GCTTTTCGCAACGATGAATGTGGCGATACATATAAAATTTGAAAGCCCGGGTACCTTACA 120
Db 58 GCTTTTCGCAACGATGAATGTGGCGGACTATATAAAATTTGAAAGCCCGGGTACCTTACA 117
QY 121 TGTCTGGTGTATCTCTCATTTATACCCCAAGTGAATAATGCGAATGGTGTATTCAGGCT 180
Db 118 TGTCTGGTGTATCTCTCATTTTACCATCCAGTGAGAAATGGAATGGTGTATTCAGGCT 177
QY 181 CCGACCCCATACAGAGAAATATGATCAACTTCAACCCCTCAGTTTCGATTGGAGGACAGA 240
Db 178 CCGAGGCCCTACAGAGAAATCATGATCAACTTCAACCCACATTTTCGATTGGAGGACAGA 237
QY 241 GACTGCAAGTATGACTACGTGGAAAGTCTTCGATGGAGAAATGAAATGGACATTTTAGG 300
Db 238 GACTGCAAGTATGACTATGTGGAAGTGTATGATGGAGAAATGAAATGGACATTTTAGG 297
QY 301 GGAAGTTCTGTGGAAGATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 298 GGAAGTTCTGTGGGAGATCGACCTTCACCTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCT 357
QY 361 ATCAAAATTTGTCTCTGACTACGAAACACATGTGTGCAAGATTTTCCATAGCTTATGAAT 420
Db 358 ATCAAAATTTGTCTCTGACTATGAGACCCACGGGCGAGGATTTTCCATCGCTATGAATC 417
QY 421 TTCAGAGAGGCTCTGAATGTTCCAGAACTACACAACTAGTGGAGTATGAAGTCC 480
Db 418 TTCAGAGAGGCGCCGGAATGTTCTCAGAACTATACAGCACCTACTGGAGTATGAAGTCC 477
QY 481 CCGGATTCCTCGAAATATCCCAACAGCTTGAATGCACCTATATGCTCTCTCTCTCTCTCT 540
Db 478 CCGGATTCCTCGAAATATCCCAACAGCTTGAATGCACCTATATGCTCTCTCTCTCTCTCT 537
QY 541 AAGATGTCAGAGATTTATCTCTGAAATTTGAAAGCTTTGACCTGGAGCTGAGTCAATCTCT 600
Db 538 AAGATGTCAGATTAATCTAGATTTTGAAGTTTGAAGCTTTGACCTGGAGCAAGACTCAATCT 597
QY 601 CCAGGGGGATGTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 560
Db 598 CCAGGGGGAATGTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 557
QY 661 GGCCTTCACATTTGGCGTTACTGTGGACAGAAACACAGGTCGAATCCGATCTCTCATCG 720
Db 658 GGCCTTCACATTTGGCGTTACTGTGGCGAGAAACTCTCTGGCGGATCCCTCTCTCTCTCTCA 717
QY 721 GGCATTTCTCTCCATGGTTTTTTTACACCGACAGCGGATAGCAAAAGAGGTTTCTCAGCA 780
Db 718 GGCATTTCTCTCCATGGTTCTTCTACACTGACAGCGCAATAGCAAGAGGTTTCTCAGCC 777
QY 781 AACTACAGTGTCTCGAGAGAGTGTCTCAGAAAGTTTCAATGATGGAAGCTCTGGGC 840
Db 778 AACTACAGCGTGTCTCGAGAGAGCATCTCTGAAGATTTTCAAGTGTATGGAGCTCTGGGC 837
QY 841 ATGGAATCAGGAGAAATTCATTTCTGACACAGATACAGCTTCTTCCAGTATAGCAACAAC 900
Db 838 ATGGAATCAGGAGAGATCCATTTCTGACACAGATACATGCAATCTTCCAGTATGATCAACA 897
QY 901 TGGTCTGACAGCGCTCCCGCTGAACTACCTGAGAAATGGGTGGAGTCCCGGAGGAGAT 960
Db 898 TGGTCTGTTGAGCGCTCCCGCTGAACTACCTGAAAGCGGTGGACACACAGGAGGAGAC 957
QY 961 TCTTACCAGAGTGGATACAGTACAGTGGGCTTCTCGGCTTCTGTACAGCTCTCGGG 1020
Db 958 TCTTACCAGAGTGGATCCAGTGGAGTGGGCTTCTCGGCTTCTGTACAGCTCTCGGG 1017
QY 1021 ACACAGGCGCCATTTCAAAAGAAACCAAGAGAAATATTTATGTCAGAGCTTACAGATC 1080
Db 1018 ACACAGGCGCCATTTCCNAGGAAACCAAGAGAAATATTTATGTCAGAGCTTACAGATC 1077

QY 1081 GACGTTAGCTCCAACGGGGAAGACTGGATCACCATAAAGAGGAAACAAACCTGTTCTC 1140
DB 1078 GACATCAGCTCCACGGAGAGGACTGGATCACCCTGAAGGAGGAAATAAAGCCATTATC 1137
QY 1141 TTTCAGGGAACACCAACCCACAGATGTTGGTTGGTTCAGTATTCGCCAAACCACTGATA 1200
DB 1138 TTTCAGGGAACACCAATCCACGGATGTTGCTTTGGAGTTTCCCCAAACCACTGATA 1197
QY 1201 ACTCGATTTGTCGGAATCAAGCCTCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1260
DB 1198 ACTCGATTTGTCGGAATCAAGCCTCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1257
QY 1261 GTATACGTTGCAAGATAACAGATTTATCTTCTCTGGAATGTTGGGTATGTTCTGGA 1320
DB 1258 GTTATGCTGCAAGATAACAGATTAACCTTACCCTTCTCTGGAATGTTGGGCATGTTCTGGA 1317
QY 1321 CTTATTTCTGACTCCCACTCACATCATCTCAACCAAGAGACAGAACTGGATGCTGAA 1380
DB 1318 CTTATTTCTGACTCCCACTCACATCATCTCAACCAAGAGACAGAACTGGATGCTGAA 1377
QY 1381 AACATCCCGCTGGTAACCAAGTCGCTTGGCTGGGCACCTTCCACCCGCACCTCATTTCCCTAC 1440
DB 1378 AACATCCCGCTGGTAACCAAGTCGCTTGGGCACCTTCCACCCGCACCTCATTTCCCTAC 1437
QY 1441 ATCAATGAGTGGCTCCAAATAGACTGGGGAGGAGAAGATCGTGAGGGGCATCATCAT 1500
DB 1438 ATCAATGAGTGGCTCCAAATAGACTGGGGAGGAGAAGATCGTGAGGGGCATCATCAT 1497
QY 1501 CAGGTGGGAAGCAGGAGACAGAAAGTGTTCATGAGGAAGTTCAAGATCGCCTACAGT 1560
DB 1498 CAGGTGGGAAGCAGGAGACAGAAAGTGTTCATGAGGAAGTTCAAGATCGCCTACAGT 1557
QY 1561 AACACCGCTCGGACTGGAAGATGATCATGGATGACACCAACGCAAGCGGAAGTCTTTT 1620
DB 1558 AACATGTTCTGACTGGAAATGATCATGGATGACACCAACGCAAGCGGAAGTCTTTT 1617
QY 1621 GAGGCAACAACTATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCCACGGGA 1680
DB 1618 GAGGCAACAACTATGATACACCTGAGCTGCGGCTTTTACACCTCTCTCCACAAAGA 1677
QY 1681 TTCATCAGGATCTACCCGAGAGAGCCACTCATGCGGACTGGGGCTCAGATGGAGCTG 1740
DB 1678 TTCATCAGGATCTACCCGAGAGAGCCACTCATAGTGGGCTCGGACTGAGGATGGAGCTA 1737
QY 1741 CTGGGCTGTGAAGTGGAAAGCCCTACAGCTGGACCGACACTCCCAACGGAAGTCTGGT 1800
DB 1738 CTGGGCTGTGAAGTGGAAAGTGGCTACAGCTGGACCGACACTCCCAACGGAAGTCTGG 1797
QY 1801 GATGAATGTGATGACGACAGGCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
DB 1798 GACGAGTGTGACGATGACAGGCCAACTGCCACAGTGGCAGGTGATGACTTCCAGCTC 1857
QY 1861 ACAGTGGCACCACCTGCTGGCCACACAAAGCCCGGCTCATAGACAGCACCATACAA 1920
DB 1858 ACAGAGGACACCTGCTGGCCACAGAGAAGCCACCATATATAGACAGCACCATACAA 1917
QY 1921 TCAGAGTTTCCAACTATGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAAGACCTTC 1980
DB 1918 TCAGAGTTTCCGACATACGTTTAACTGCGAGTTTGGCTGGGGCTCTCACAAAGACATTC 1977
QY 1981 TGCCACTGGGAACATGACAATCAGCTGCAGCTCAAGTGGAGTGTGTTGACCAGCAAGACG 2040
DB 1978 TGCCACTGGGAACATGACAAGCACGCGCAGCTCAGTGGAGGGTCTGACCAAGCAAGCG 2037
QY 2041 GGACCATTCAGGATCACACAGGAGATGGCAACTTCATCTATTCCCAAGCTGACCAAAAT 2100
DB 2038 GGCCCATTCAGGACCAACAGGAGATGGCAACTTCATCTATTCCCAAGCTGATGAAAT 2097
QY 2101 CAGAAAGGCAAGTGGCTGCTGCTGAGCCCTGTTGGTTTATTTCCAGAACTCTGCCAC 2160
DB 2098 CAGAAAGGCAAGTGGCTGCTGCTGAGCCCTGTTGGTCTATTTCAGAGTTCTGCCAC 2157
QY 2161 TGCATGACCTTCTGGTATCACATGTCTGGGTCCCAAGCTCGGCACACTCAGGGTCAAACCTG 2220

DB 2158 TGCATGACCTTCTGGTATCACATGTCCGCTCTCATGTGGGTACACTGAGGCTCAAACCTG 2217
QY 2221 CGCTACCAGAGCCAGAGGAGTACGATCAGCTGGTCTGGATGGCCATTTGGACACCAAGT 2280
DB 2218 CACTACCAGAGCCAGAGGAATATGATCAACTGGTCTGGATGGTCTGGGACCAAGGA 2277
QY 2281 GACCACTGGAAGGAGGCGTCTTGTCTCCACAAGTCTCTGAAACTTTTATCAGCTGATT 2340
DB 2278 GACCACTGGAAGGAGGCGTCTTGTCTGCACAATCTCTGAAACTGTATCAGGTTATT 2337
QY 2341 TTCGAGGCGAAATCGGAAAAGAACCTTGGTGGGATGCTGTGGATGACATTTAGTATT 2400
DB 2338 TTTGAAGGTGAAATCGGAAAAGGAACCTCGGTGGGATTTGCTGTGGATGATATCAGTATT 2397
QY 2401 AATAACCACATTTCAACAAGAAGATTGTCAAAACACAGACCTCTGGATAAAAGAACCCA 2460
DB 2398 AACAAACCACATTCCTTCAGGAGGACTGTCAAAACCAACAGACCTAGATAAAAAGAACACA 2457
QY 2461 GAAATTAATAATTTGATGAACAGGAGCAGCCAGGATACGAAGGTGAAGGAGAACGCTGAC 2520
DB 2458 GAAATTAATAATGATGAACAGGAGCAGCCAGGATATGAA--GAAGGGAAGGCGAC 2514
QY 2521 AAGAACATCTCAGGAAGCCAGGCAATGTGTTGAAGACCTTTAGAACCATCCTCATCACC 2580
DB 2515 AAGAACATCTCAGGAAGCCAGGCAATGTGCTTAAAGACCTGGACCCCATCCTGATCACC 2574
QY 2581 ATCATAGCCATCAGCGCCCTGGGGTCTCTCTGGGGCTGTCTGGGGTCTGTGGTGTAC 2640
DB 2575 ATCATAGCCATCAGTGGCCCTGGGGTCTCTCTGGGTGCTGTGGGATTTGCTGTGTAC 2634
QY 2641 TGTGCTCTTGGCATTAATGGGATGTCAAGAAAGAACTTTGTCTGCCCTGGAGAACTATAAC 2700
DB 2635 TGTGCTCTTGGCACAATGGGATGTGCGAAGGAACCTATCTGCCCTGGAGAACTATAAC 2694
QY 2701 TTTGAACCTTGTGGATGTTGTTGAAAGTTGAAAAAGAACAACTGAATACACAGAGTACTTAT 2760
DB 2695 TTTGAACCTTGTGGATGTTGTTGAAAGTTGAAAAAGATAAACTGAACCCACACAGTAATTAC 2754
QY 2761 TCGGAGGCGATGA 2772
DB 2755 TCAGAGGCGTGA 2766
RESULT 15
AAZ31462 standard; cDNA; 2766 BP.
ID AAZ31462 standard; cDNA; 2766 BP.
AC AAZ31462;
AT 07-FEB-2000 (first entry)
DE Rat neuropilin cDNA sequence.
EX Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; rat; ss.
OS Rattus sp.
XX WO955855-A2.
FN 04-NOV-1999.
XX 23-APR-1999; 99WO-CA00324.
PR 23-APR-1998; 98US-0082791.
XX (GENE-) GENESENSE TECHNOLOGIES INC.
PI Wright JA, Young AH, Lee YS;
XX WPI; 2000-023357/02.

PT Antisense oligonucleotides that inhibit neuropilin expression, useful
XX for treating cancer -
PS Disclosure; Fig 6; 57pp; English.
XX
CC The invention provides sequences AAZ31431-460 which are antisense
CC oligonucleotides that inhibit human neuropilin expression. The antisense
CC oligonucleotides can be used to inhibit the growth or metastasis of a
CC mammalian tumor and inhibit neovascularisation. The oligonucleotides may
CC be used to treat various forms of cancers or tumors, such as sarcomas,
CC melanomas, adenomas, carcinomas of solid tissue, hypoxic tumors,
CC squamous cell carcinomas of the mouth, throat, larynx and lung,
CC genitourinary cancers such as cervical and bladder cancer, hematopoietic
CC cancers, colon cancer, breast cancer, pancreatic cancer, renal cancer,
CC brain cancer, skin cancer, liver cancer, head and neck cancers, and
CC nervous system cancers, as well as benign lesions such as papillomas. The
CC methods may be used to treat neovascularisation disorders such as
CC diabetic retinopathy, and retinopathy of prematurity and age related
CC macular degeneration. The present sequence represents the rat neuropilin
CC cDNA.
XX
SQ Sequence 2766 BP; 766 A; 680 C; 716 G; 604 T; 0 other;

Query Match 79.2%; Score 2196; DB 21; Length 2766;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 2426; Conservative 0; Mismatches 340; Indels 6; Gaps 2;

QY 1 ATGGAGAGGGGGTGGCGTCTCTCGCGGTGCTCGCCCTCGCTCGCCCGCGCGG 60
DB 1 ATGGAGAGGGGGTGGCGTCTCTCGCGGTGCTCGCCCTCGCTCGCCCGCGG 57
QY 61 GCTTTTCGCAACGATGAATGTGGCGTACTATATAAAATTGAAAGCCCGGGTACCTTACA 120
DB 58 GCTTTTCGCGAGCATAAATGTGGCGGACTATATAAAATTGAAAGCCCGGGTACCTTACA 117
QY 121 TCTCCTGGTTATCTCTTATATACCAAGTGAATGGAATGCGATTCAGGCT 180
DB 118 TCTCCCGGTACCTCATTTCTTACCATTCAAGTGAATGGAATGCGATTCAGGCT 177
QY 181 CCGAGCCCATACAGAGAATATGATCAACTTCAACCCCTCACTTCGATTTGGAGGACAGA 240
DB 178 CCGAGCCCTACAGAGATCATGATCACTTCAACCCACATTTCCATTTGGAGGACAGA 237
QY 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAATGAAATGGACATTTAGG 300
DB 238 GACTGCAAGTATGACTATGTGAAGTATCGATGGAGAAATGAAGTGGCCGCTGTGG 297
QY 301 GGAACTTCTGTGGAAGATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTTT 360
DB 298 GGAAGTTCTGTGGAAGATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTC 357
QY 361 ATCAAAATTTGCTCTGACTACGAAACACATGGTGCAGGATTTTCCATAGTTTGAATTT 420
DB 358 ATCAAAATTTGCTCTGACTATGAGACCCAGGGGAGGATTTTCCATCCGCTATGAATC 417
QY 421 TTCAAGAGAGTCTGGAATTTTCCAGAACTACACACACCTAGTGGAGTGAATAAGTCC 480
DB 418 TTCAAGAGAGGGCCGGAATTTCTCAGAACTATACAGCACCTACTGGAGTGAATAAGTCC 477
QY 481 CCGGGATTCCTGAAATAATCCCAACAGCCTTGAATGCACTTATTTCTTTGGGCCA 540
DB 478 CTGGGTTCCCTGAAATAATACCCCAACAGCTTGGAGTGCACCTACATCTTTGCAACA 537
QY 541 AAGATCTCAGAGATATCTCGGAATTTGAAAGCTTTGACCTGGAGCCTCACTCAATCCT 600
DB 538 AAGATCTCAGAGATATCTCAGATTTGAAAGTTTGAAGTTTGAAGTGCAGCAACTCAATCCT 597
QY 601 CCAGGGGGGATGTTCTGCTACGACCGGCTAGAAATCTGGGATGGATTCCTGATGTT 660
DB 598 CCGGAGGAATGTTCTGCTGCTATGACCGGCTGGAGATCTGGAGTGGATTCCTGAAGTT 657
QY 661 GGCCTCACAATTTGGGCTTACTGTGGACAGAAACACAGGTCGAATCCGATCCTCATCG 720
DB 1738 CTGGGCTGTGAAGTGAAGTGCCTACAGTGGACCCACACACCCCAATGGGAACCCCGTG 1799

DB 658 GGCCCTCACATTTGGCGTCTACTGTGGCGAGAAACATCTCTGGCGGATCCCTCTCTTCA 717
QY 721 GGCATTTCTCTCCATGGTTTTTTTACACGACGCGCATAGACAAAGAGTTTCTCTCAGCA 780
DB 718 GGCATTTCTATCCATGGTCTTCTACTGACAGCGCAATAGCAAGNAGGTTTCTCAGCC 777
QY 781 AACTACAGTGTCTTTGAGAGCAGTGTCTCAGAAAGATTTCAAATGTATGGAAGTCTGGGC 840
DB 778 AACTACAGGCTGTCTGAGAGCAGCATCTCTGAAGATTTCAAGTGTATGGAGGCTCTGGGC 837
QY 841 ATGGAATCAGGAGAAATTCATCTGACCAAGATCAGAGCTTCTTCCAGTATACCAACAAC 900
DB 838 ATGGAATCTGGAGAGATCCATTTCTGACCAAGATCAGTCATCTTCCCAGTATGTTACCAAC 897
QY 901 TGGTCTGACAGAGCGTCCCGCTTAACTACCTGAGAAAGTGGTGGACTCCCGGAGAGGAT 960
DB 898 TGGTCTGTTGAGCGTCCCGCTTAACTACCTGAGAAAGTGGTGGAGCAGAGAGGAG 957
QY 961 TCTTACCGAGAGTGGATACAGTGTGGCTTCTTGGCTTCTGCGCTTTGTCAGCGGCTGCGG 1020
DB 958 TCCCTACAGGAGTGGATCCAGGTGGACTTGGGCTCTCTGCGATTCGTTACTGCTGTGGG 1017
QY 1021 ACACAGGGGCCATTTCAAAGAAACCAAGAGAAATATATGTCAGAGCTTACAAGATC 1080
DB 1018 ACACAGGGTGCATTTTCCAAGGAAACCAAGAGAAATATATGTCAGAGCTTACAAGATA 1077
QY 1081 GAGCTTAGCTCCAAACGGGAGAGCTGATCACCATAAAGAGGAAACAAACCTGTTCTC 1140
DB 1078 GACATCAGTCTCCAAACGGAGAGGACTGATCACCCTGAGAGGAGGAAATGAAGCCATATC 1137
QY 1141 TTTTACGGGAAACACCAACCCACAGATTTGTGTGGTGTGTCAGTATTTCCCAACCACTGATA 1200
DB 1138 TTTTACGGGAAACACCAATCCACGGATTTGTCTTTGGAGTTTCCCAACCACTGATA 1197
QY 1201 ACTCGATTTGTCCGATCAAGCTGCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1260
DB 1198 ACTCGATTTGTCCGATCAAACTGCACTCTGGGAACTGGAAATCTATGAGATTTGAA 1257
QY 1261 GTATACGTTGCAAGATAACAGATTTATCTGCTCTGGAATTTGGTATTTGGTCTGGA 1320
DB 1258 GTTATGGCTGCAGATACAGATTTACCTTGTCTTGGATTTGGGCATTTGGTCTGGA 1317
QY 1321 CTTATTTCTGACTCCAGATCACATCATCCAAACAGGAGACAGAACTGGATGTCCTGAA 1380
DB 1318 CTTATTTTCAAGTCTCCAGATTTACAGCATCCAAACAGGAGACAGAACTGGATGCCAGAA 1377
QY 1381 AACATCCGCTGGTACCAGTCTGCTGCTGGGCACTTCCACCCGCACTCATTTCTTAC 1440
DB 1378 AACATCCGCTGGTACCAGTCTGCAACCGGCTGGGCTGCCACCTCACCCCACTATAC 1437
QY 1441 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAGAACTGTCAGGGGCACTCATCT 1500
DB 1438 ATCAATGATGCTCCAACTGGAGCTGGAGATGAGAACTAGTAGAGGTGTCATCTT 1497
QY 1501 CAGGTTGGGAAGCACCAGAGAACAAAGTGTTCATGAGAAAGTTCAAGATTCGGGTACAG 1560
DB 1498 CAAAGTTGGGAAGCACCAGAGAAACAAAGTGTTCATGAGAAAGTTCAAGATTCGCTACAGT 1557
QY 1561 AACACGGCTCCGACTGGGAAGATGATCATGATGACAGCAACAGCAAGCGGAAGTCTTTT 1620
DB 1558 AACAAATGTTCTGACTGGAAATGATCATGATGACAGCAAGCGCAAGCTAAGTCTTTT 1617
QY 1621 GAGGCAACAACTATGATACCTGAGCTGCGGACTTTTCCAGCTCTCTCCAGCGCA 1680
DB 1618 GAAGCAACAACTATGACACACTGAGCTCCGGGCTTTTACACCTCTCTCCACAGA 1677
QY 1681 TTCATCAGGATCTACCCGAGAGAGCCACTCATGCGGACTGGGCTCAGATTCAGATTCGAGCTG 1740
DB 1678 TTCATCAGGATCTACCCGAGAGAGCCACATAGTGGGCTCGGACTGAGGATGGAGCTA 1737
QY 1741 CTGGGCTGTGAAGTGAAGCCCTTACAGCTGGACCGACCACTCCCAACCGGGAAGTCTGGTG 1800
DB 1738 CTGGGCTGTGAAGTGAAGTGCCTACAGTGGACCCACACACCCCAATGGGAACCCCGTG 1799

THIS PAGE BLANK (USPTO)

QY 61 GCTTTTCGCAACGATGAATGTGGGATACATATAAATAATGAAAGCCCGGGTACCTTTACA 120
DB 61 GCTTTTCGCAACGATGAATGTGGGATACATATAAATAATGAAAGCCCGGGTACCTTTACA 120
QY 121 TCTCCTGGTTATCCTCATCTCTTATCACCAGTGAATAATGCGAATGGCTGATTCAGGCT 180
DB 121 TCTCCTGGTTATCCTCATCTCTTATCACCAGTGAATAATGCGAATGGCTGATTCAGGCT 180
QY 181 CCGGACCCATPACGAGAAATATGATCACTTCAACCCCTCACCTTCGATTTGGAGGACAGA 240
DB 181 CCGGACCCATPACGAGAAATATGATCACTTCAACCCCTCACCTTCGATTTGGAGGACAGA 240
QY 241 GACTGCAAGTATGACTACGTTGGAAGTCTTCGATGGAGAAATGAAAATGACATTTTAGG 300
DB 241 GACTGCAAGTATGACTACGTTGGAAGTCTTCGATGGAGAAATGAAAATGACATTTTAGG 300
QY 301 GGAAGTCTCTGGAAGATAGCCCTCTCCTCTGTTGTGTCTTCAGGGCCATTTCTTTT 360
DB 301 GGAAGTCTCTGGAAGATAGCCCTCTCCTCTGTTGTGTCTTCAGGGCCATTTCTTTT 360
QY 361 ATCAAAATTTCTCTGACTACGAAACACATGTTGCGAGGATTTTCCATACGTTATGAAAT 420
DB 361 ATCAAAATTTCTCTGACTACGAAACACATGTTGCGAGGATTTTCCATACGTTATGAAAT 420
QY 421 TTCAAGAGAGTCTCTGAATGTTCCCAAGACTACACACACCTAGTGGAGTGATAAGTCC 480
DB 421 TTCAAGAGAGTCTCTGAATGTTCCCAAGACTACACACACCTAGTGGAGTGATAAGTCC 480
QY 481 CCGGGATTCCTGAAATAATCCACAGCCTTGAATGACATTAATGCTTTTGGGCCA 540
DB 481 CCGGGATTCCTGAAATAATCCCAAGCCTTGAATGACATTAATGCTTTTGGGCCA 540
QY 541 AAGATCTCAGAGATTAATCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAAATCCT 600
DB 541 AAGATCTCAGAGATTAATCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAAATCCT 600
QY 601 CAGAGGGGATGTTCTGCTGCTACGACCGGCTAGAAATCTGGGATGGATTCCTGATGTT 660
DB 601 CAGAGGGGATGTTCTGCTGCTACGACCGGCTAGAAATCTGGGATGGATTCCTGATGTT 660
QY 661 GGCCTCTACATTTGGGGTACTGTGACAGAAACACAGGTCGAATCCGATCCTCATCG 720
DB 661 GGCCTCTACATTTGGGGTACTGTGACAGAAACACAGGTCGAATCCGATCCTCATCG 720
QY 721 GGCATCTCTCCATGTTTTTACACCGACAGCGGATAGCAAAAGAGTTTCTCAGCA 780
DB 721 GGCATCTCTCCATGTTTTTACACCGACAGCGGATAGCAAAAGAGTTTCTCAGCA 780
QY 781 AACTACAGTGTCTTGACAGCAGTGTCTCAGAAAGATTTCAAATGTATGGAAGCTCTGGC 840
DB 781 AACTACAGTGTCTTGACAGCAGTGTCTCAGAAAGATTTCAAATGTATGGAAGCTCTGGC 840
QY 841 ATGGAATCAGGAGAAATTCATTTGTACACAGATCACAGCTTCTCCAGTATAGCAACAAC 900
DB 841 ATGGAATCAGGAGAAATTCATTTGTACACAGATCACAGCTTCTCCAGTATAGCAACAAC 900
QY 901 TGGTCTGCAGAGCGCTCCCGCTGAACTACCCCTGAGAATGGGTGGACTCCCGAGAGGAT 960
DB 901 TGGTCTGCAGAGCGCTCCCGCTGAACTACCCCTGAGAATGGGTGGACTCCCGAGAGGAT 960
QY 961 TCTACCGAGATGGATACAGTAGACTTTGGGCTTCTCGGCTTTGTCTACGGGTCTGGG 1020
DB 961 TCTACCGAGATGGATACAGTAGACTTTGGGCTTCTCGGCTTTGTCTACGGGTCTGGG 1020
QY 1021 ACACAGGGCCATTTCAAAGAAACCAAGAAATATATGTCAAGACTTACAGATC 1080
DB 1021 ACACAGGGCCATTTCAAAGAAACCAAGAAATATATGTCAAGACTTACAGATC 1080
QY 1081 GACGTTAGTCTCAACGGGGAAGACTGGATCACCATAAAGAGAAACAAACCTGTTCTC 1140
DB 1081 GACGTTAGTCTCAACGGGGAAGACTGGATCACCATAAAGAGAAACAAACCTGTTCTC 1140
QY 1141 TTTACAGGAAACCAACCCACAGATGTTGGTTGCGATATTCGCCAAACCACTGATA 1200

DB 1141 TTTACAGGAAACCAACCCACAGATGTTGGTTGCTAGTATTTCCCAACCACTGATA 1200
QY 1201 ACTCGAATTTCTCGAATCAAGCTTGAACCTTGGAACTGCGCATATCTATGAGATTTGAA 1260
DB 1201 ACTCGAATTTCTCGAATCAAGCTTGAACCTTGGAACTGCGCATATCTATGAGATTTGAA 1260
QY 1261 GTATAGCGTTGCAAGATAACAGATTAATCCTTGTCTGGAATGTTGGGTATGGTGTCTGGA 1320
DB 1261 GTATAGCGTTGCAAGATAACAGATTAATCCTTGTCTGGAATGTTGGGTATGGTGTCTGGA 1320
QY 1321 CTTATTTCTGACTCCGAGATACATCATCAACCAAGGAGACAGAACTGGATGCTGTGAA 1380
DB 1321 CTTATTTCTGACTCCGAGATACATCATCAACCAAGGAGACAGAACTGGATGCTGTGAA 1380
QY 1381 AAGATCGGCTGTGAACAGTCTGCTGCTGGCTGGGACTTCCACCCGACCTCATCTCTAC 1440
DB 1381 AAGATCGGCTGTGAACAGTCTGCTGCTGGCTGGGACTTCCACCCGACCTCATCTCTAC 1440
QY 1441 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAAAGATCGTGAGGGGCATCATCAT 1500
DB 1441 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAAAGATCGTGAGGGGCATCATCAT 1500
QY 1501 CAGGTTGGGAAGCACCGAGAGAAACAGGTTTCTATGAGGAAGTTCAAGATCGGGTACAGC 1560
DB 1501 CAGGTTGGGAAGCACCGAGAGAAACAGGTTTCTATGAGGAAGTTCAAGATCGGGTACAGC 1560
QY 1561 AACAACGGCTCGGACTGGAAGATGATCATGATGACAGCAACCAAGGCGAAGTCTTTT 1620
DB 1561 AACAACGGCTCGGACTGGAAGATGATCATGATGACAGCAACCAAGGCGAAGTCTTTT 1620
QY 1621 GAGGCAACAACAACTATGATACACCTGAGCTCGGACTTTTCCAGCTCTCTCCACGCGA 1680
DB 1621 GAGGCAACAACAACTATGATACACCTGAGCTCGGACTTTTCCAGCTCTCTCCACGCGA 1680
QY 1681 TTCAATCAGATCTACCCCGAGAGAGCCATCATGCGGACTGGGGTCTCAGATGAGAGCTG 1740
DB 1681 TTCAATCAGATCTACCCCGAGAGAGCCATCATGCGGACTGGGGTCTCAGATGAGAGCTG 1740
QY 1741 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACCCACCACTCCCAACGGGAAGTGGTG 1800
DB 1741 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACCCACCACTCCCAACGGGAAGTGGTG 1800
QY 1801 GATGAATGTGATGACGACGAGCCAACTGCCAGTGGAAACAGGTGATGACTTCCAGCTC 1860
DB 1801 GATGAATGTGATGACGACGAGCCAACTGCCAGTGGAAACAGGTGATGACTTCCAGCTC 1860
QY 1861 ACAGGTGGCAACCACTGTGTGGCCACAGAAACCCACGGTATAGACAGCACCATTACAA 1920
DB 1861 ACAGGTGGCAACCACTGTGTGGCCACAGAAACCCACGGTATAGACAGCACCATTACAA 1920
QY 1921 TCAGAGTTTCCAACTATGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAGACCTTC 1980
DB 1921 TCAGAGTTTCCAACTATGTTTAACTGTGAATTTGGCTGGGGCTCTCACAAGACCTTC 1980
QY 1981 TGCCACTGGGAACATGACAATCACGTCGAGCTCAAGTGGAGTGTGTGACCAGCAAGAG 2040
DB 1981 TGCCACTGGGAACATGACAATCACGTCGAGCTCAAGTGGAGTGTGTGACCAGCAAGAG 2040
QY 2041 GGACCCATTCAGATCACACAGAGATGGCAACTTCATCTATTTCCCAAGCTGACGAAAT 2100
DB 2041 GGACCCATTCAGATCACACAGAGATGGCAACTTCATCTATTTCCCAAGCTGACGAAAT 2100
QY 2101 CAGAAGGGCAAGTGGCTCGCTGGTGGCCCTGTGTTTATTTCCAGAACTCTGCCAC 2160
DB 2101 CAGAAGGGCAAGTGGCTCGCTGGTGGCCCTGTGTTTATTTCCAGAACTCTGCCAC 2160
QY 2161 TGCATGACCTTCTGCTATCACATGCTGGGTCCCGCTGGGACACTCAGGGTCAAACTG 2220
DB 2161 TGCATGACCTTCTGCTATCACATGCTGGGTCCCGCTGGGACACTCAGGGTCAAACTG 2220
QY 2221 CGCTACCAAGCCAGAGAGGTACGATCAGTGGTGTGATGGCCATTTGGACCAACAGGT 2280

Db 2221 CGTACCAGAGCCAGAGGAGTACGATCAGCTGCTCTGGATGGCCATTGGACACCAAGT 2280
Qy 2281 GACCACTGGAAGGAGCGGTCTCTCCACAGTCTCTGAAACTTATCAGGTGATT 2340
Db 2281 GACCACTGGAAGGAGCGGTCTCTCCACAGTCTCTGAAACTTATCAGGTGATT 2340
Qy 2341 TTCGAGGCGGAATCGGAAAAAGAACCTTGGTGGATTGCTGTGGATGACATTAGTATT 2400
Db 2341 TTCGAGGCGGAATCGGAAAAAGAACCTTGGTGGATTGCTGTGGATGACATTAGTATT 2400
Qy 2401 AATAACCCACATTTACACAGAAGATTGTCCAAACCCAGACGCTGGATGAAAAAGAACCA 2460
Db 2401 AATAACCCACATTTACACAGAAGATTGTCCAAACCCAGACGCTGGATGAAAAAGAACCA 2460
Qy 2461 GAAATTAATAATTTGATGAACAGGAGCACGCCAGGATACGAAGGTGAAGGAGAGGTGAC 2520
Db 2461 GAAATTAATAATTTGATGAACAGGAGCACGCCAGGATACGAAGGTGAAGGAGAGGTGAC 2520
Qy 2521 AAGAACATCTCCAGGAAGCCAGGCAATGTGTGAAGACCTTAGAACCCATCCTCATCACC 2580
Db 2521 AAGAACATCTCCAGGAAGCCAGGCAATGTGTGAAGACCTTAGAACCCATCCTCATCACC 2580
Qy 2581 ATCATAGCCATGAGCGCCCTGGGGTCTCTCTGGGGGTCTGTGTGGGGTCTGTGTGTAT 2640
Db 2581 ATCATAGCCATGAGCGCCCTGGGGTCTCTCTGGGGGTCTGTGTGGGGTCTGTGTGTAT 2640
Qy 2641 TGTGCTCTGTGGCATAATGGGATGTCAGAAAGAACTTGTCTGCCCTGGAGAACTATAC 2700
Db 2641 TGTGCTCTGTGGCATAATGGGATGTCAGAAAGAACTTGTCTGCCCTGGAGAACTATAC 2700
Qy 2701 TTTGAACCTTGTGGATGGTGTGAAGTTGAAAAAGACAACTGAATACACAGAGTACTTAT 2760
Db 2701 TTTGAACCTTGTGGATGGTGTGAAGTTGAAAAAGACAACTGAATACACAGAGTACTTAT 2760
Qy 2761 TCGGAGGCATGA 2772
Db 2761 TCGGAGGCATGA 2772

RESULT 2

US-09-116-473-3
; Sequence 3, Application US/09116473
; Patent No. 6428965
; GENERAL INFORMATION:
; APPLICANT: Kolodkin, Alex
; APPLICANT: Ginty, David
; TITLE OF INVENTION: SEMAPHORIN RECEPTOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/116.473
; FILING DATE: 17-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/052,762
; FILING DATE: 17-JUL-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32141
; REFERENCE/DOCKET NUMBER: 01107.74973
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100

; TELEFAX: 202-508-9299
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3471 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-116-473-3

Query Match 79.7%; Score 2210.4; DB 4; Length 3471;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 2428; Conservative 0; Mismatches 341; Indels 3; Gaps 1;

Qy 1 ATGAGAGGGGCTGCCCTCTCTGCGCGGTGCTGCGCCCTGCTCTGCGCCGGGTACCTTACA 120
Db 145 ATGAGAGGGGCTGCCCTCTCTGCGCGCGGTGCTGCGCCCTGCTCTGCGCCGGGTACCTTACA 204
Qy 61 GCTTTTCGAACGATGAATGTGGCGATATATAAAATTTGAAAGCCCGGGGTACCTTACA 120
Db 205 GCTTTTCGAACGATGAATGTGGCGGACTATATAAAATTTGAAAGCCCGGGGTACCTTACA 264
Qy 121 TCTCTGTTTATCTTCATCTTATCACCAGTGAATAATGCGAATGCGTGAATCAGGCT 180
Db 265 TCTCCGGCTACCCCTCATCTTACCATCAAAGTGAGAAATGTGAATGGCTAATCCAAGCT 324
Qy 181 CCGAGCCCATACCAGAGAAATATGATCAACTCAACCCTCACTTCGATTTGGAGGACAGA 240
Db 325 CCGAGCCCATACCAGAGAAATATGATCAACTCAACCCTCACTTCGATTTGGAGGACAGA 384
Qy 241 GACTGCAAGTATGACTACGTGGAAGTCTTCGATGAGAGAAATGAAATGGACATTTTAGG 300
Db 395 GACTGCAAGTATGACTATGTGGAAGTATGATGATGATGAGAGAAATGAAAGTGGCGCTGTGG 444
Qy 301 GGAAGTCTGTGGAAGATAGCCCTCTCTCTGTGTGTCTTCAGGGCCATTTCTTTT 360
Db 445 GGAAGTCTGTGGAAGATAGCCCTCTCTCTGTGTGTCTTCAGGGCCATTTCTTTT 504
Qy 361 ATCAATTTGCTCTGACTACGAAACACATGTCGAGAGATTTTCCATACGTTATCAAAAT 420
Db 505 ATCAATTTGCTCTGACTATGAGACCCACGCGGCGAGAGATTTTCCATCGGCTATGAATC 564
Qy 421 TTCAGAGAGGTCTCTGAATGTTCGAGAACTACACACACCTAGTGGAGTGAATAAGTCC 480
Db 565 TTCAGAGAGGTCTCTGAATGTTCGAGAACTATACAGACCTACTGAGTGAATAAGTCC 624
Qy 481 CCGGATTCCTGAAAAATATCCCAACAGCCTTGAATGCACTTATATGCTTTTCGGCA 540
Db 625 CCGGATTCCTGAAAAATATCCCAACAGCCTTGAATGCACTTATATGCTTTTCGGCA 684
Qy 541 AAGATGTCAGAGATTTATCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAAATCCT 600
Db 685 AAGATGTCAGAGATTTATCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAAATCCT 744
Qy 601 CCAGGGGATGTTCTGCTGCTACGACCGGCTAGAAATCTGGGATGGATTCCTGATGTT 660
Db 745 CCAGGGGATGTTCTGCTGCTATGACCGGCTGGAGATCTGGATGGATTCCTGATGTT 804
Qy 661 GGCCCTCACATGGGCGTTACTGTGGACAGAAAAACACAGGTGCAATCCGATTCATCG 720
Db 805 GGCCCTCACATGGGCGTTACTGTGGGACAGAAAACTCCTGGCGGATCCGCTCTCTTCA 864
Qy 721 GGCATTCCTCATGGTGTGTTTACACCGCAGACGCGGATGAGAAAGAGAGTTTCTCAGCA 780
Db 865 GGCATTCCTCATGGTGTGTTTACACCGCAGACGCGGATGAGAAAGAGAGTTTCTCAGCA 924
Qy 781 AACTACAGTGTCTGCAGAGAGTGTCTCAGAAAGATTTCAAATGTATGAAAGCTCTGGCC 840
Db 925 AACTACAGTGTCTGCAGAGAGTGTCTCAGAAAGATTTCAAATGTATGAAAGCTCTGGCC 984
Qy 841 ATGGAATCAGGAGAAATTCATTTGACAGATTCACAGCTTCTTCCAGTATAGCACCAAC 900
Db 985 ATGGAATCAGGAGATTCATTTGACAGATTCACAGCTTCTTCCAGTATAGCACCAAC 1044

ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936.135
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC97-288-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3652 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-936-135-5

Query Match 79.5%; Score 2202.4; DB 3; Length 3652;
Best Local Similarity 87.2%; Pred. No. 0;
Matches 2416; Conservative 0; Mismatches 356; Indels 0; Gaps 0;

QY 1 ATGGAGAGGGGCTCCGCTCCTCTGCGCGGTGCTCGCCCTGCTCGCCCGCGCGGC 60
DB 348 ATGGAGAGGGGCTCCGCTGCTGTGTGCGCCAGCTCGCCCTGCGCCCTGCGCGGC 407
QY 61 GCTTTTCGCAAGATGATGGCGATGACTATAAATAATGAAAGCCCGCGGTACCTTACA 120
DB 408 GCTTTTCGCGAGCAATGTCGCGGAGCAATAAATAATGAAAGCCCGCGGTACCTTACA 467
QY 121 TCTCTGGTATCTCTCATCTTATCAACCAAGTGAATAATGGAATGGCTGATTCAGGCT 180
DB 468 TCTCCGGTTACCTCATCTTACCATCCAGTGAGAGTGTGAATGGTAATCCAGCT 527
QY 181 CCGGACCCATACGAGAAATATGATCAACTTCAACCCCTCACTTCGATTTGGAGGACAGA 240
DB 528 CCGGAAACCTACGAGAAATCAATACTCAACCCCACTTCGATTTGGAGGACAGA 587
QY 241 GACTCAAGTATGACTAGCTGGAAGTCTTCGATGAGAGAAATGAAATGGACATTTTAGG 300
DB 588 GACTCAAGTATGACTAGCTGGAAGTAAATGATGGGAGAAATGAAAGCGCGCCCTGTGG 647
QY 301 GGAAGTTCTGTGGAAGATAGCCCTCTCTCTGTTGTGTTCTTCAAGGCAATTTCTTTT 360
DB 648 GGAAGTTCTGTGGAAGATAGCCCTCTCTGTTGTGTTCTTCAAGGCAATTTCTTTT 707
QY 361 ATCAAAATTTGCTGTGACTACGAAACACATGCTGAGGATTTCCATACGTTATGAAAT 420
DB 708 ATCAAAATTTGCTGTGACTATGAGACACATGGCGAGGGGTTTTCATCCGCTATGAAATC 767
QY 421 TTCAAGAGGTCCTGAATGTTCCAGAACTACACAACTAGTGGAGTGATGAAGTCC 480
DB 768 TTCAAGAGGTCCTGAATGTTTCTCAGAACTATACAGCACTTACTTGGAGTGAAGTCC 827
QY 481 CCCGATTCCCTGAAATATATCCCAACAGCCTTGAATGACATATATGCTTTTCGGCA 540
DB 828 CTGGTTCCCTTGAATAATACCCCAACTGCTTGGAGTGACCTATCATCTTTGCACCA 887
QY 541 AAGATGTCAGAGATTTCTTGGAAATTTGAAAGCTTTGACCTGGAGCCTGACTTCAATCT 600
DB 888 AAGATGTCAGAGATTTCTTGGAGTTTGAAGCTTTTGAAGCTTTGACCTGGAGCCTGACTTCAATCT 947
QY 601 CCAGGGGAGATTTCTGTGCTACGACCGGCTAGAAATCTGGGATGGATTCCTGATGTT 660
DB 948 CCCGAGGAATGTTCTGTGCTATGACCGGCTTGGAGATCTGGGATGGATTCCTGAAAT 1007

QY 661 GGCCCTCACATTGGGCGTTACTGTGACAGAAAAACACAGGTGCGAATCCGATCCTCATCG 720
DB 1008 GGCCCTCACATTGGGCGTTACTGTGCGCAGAAAAACTCTGCGCGGATCGGCTCTCTTCA 1067
QY 721 GGCATTCTCTCCATGGTTTTTTTACACGACAGCGGATAGCAAAAAGAGTTTCTCAGCA 780
DB 1068 GGCATTCTCTCCATGGTTTTTTTACACTGACAGCGCAATAGCAAAAAGAGTTTCTCAGCC 1127
QY 781 RACTACAGTGTCTTCGACAGCAGTGTCTCAGAAAGATTTCAAAATGTATGAAAGCTCTGGC 840
DB 1128 AACTACAGTGTCTTCGACAGCAGTGTCTCAGAAAGATTTTAAAGTGTATGGAGCTCTGGC 1187
QY 841 ATGGAATCAGGAGAAATTCATTCTCACCAGATCACAGCTTCTTCCAGTATAGCACCAAC 900
DB 1188 ATGGAATCTGGAGAGATCCATCTCTCATCATCATCTGATCTTCACAGTATGGTACCAAC 1247
QY 901 TGGTCTGACAGCGCTCCCGCTGAACCTGAGATGAGATGGGTGGAGTCCCGGAGAGGAT 960
DB 1248 TGGTCTGACAGCGCTCCCGCTGAACCTGAGATGAGATGGGTGGAGTCCCGGAGAGGAT 1307
QY 961 TCCTACCGAGAGTGTGATACAGTGTGAGATGAGATGGGTGGAGTCCCGGAGAGGAT 1020
DB 1308 TCCTACCGAGAGTGTGATACAGTGTGAGATGAGATGGGTGGAGTCCCGGAGAGGAT 1367
QY 1021 ACACAGGCGGCTTCAAAAGAAAAACCAAGAAATATATGTCAAGACTTACAAAGATC 1080
DB 1368 ACACAGGCTGCTTCCAGGAAACCAAGAAATATATGTCAAGACTTACAAAGATC 1427
QY 1081 GAGCTTAGCTCCACGGGAGAGTGGATCACTACCATAAAGAGAAACAAACCTGTTCTC 1140
DB 1428 GACATCAGCTCCACGGGAGAGTGGATCACTACCATAAAGAGAAATAAAGCCATATC 1487
QY 1141 TTTGAGGAAACACAAACCCACAGATGTTGGTGTGAGTATTTCCCAACCACTGATA 1200
DB 1488 TTTGAGGAAACACAAACCCACAGATGTTGGTGTGAGTATTTCCCAACCACTGATA 1547
QY 1201 ACTCATTTGTCGGAATCAAGCTTCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1260
DB 1548 ACTCATTTGTCGGAATCAAGCTTCAACTTGGGAACTGGCATATCTATGAGATTTGAA 1607
QY 1261 GPATACGTTGCAAGATAACAGATTTATCTGCTCTGGAATTTGGGTATGGTGTCTGGA 1320
DB 1608 GTTATGCTGCAAGATAACAGATTTATCTGCTCTGGAATTTGGGATGTTGTTCTGGA 1667
QY 1321 CTTATTTCTGACTCCCACTACATCATCAACCAAGGAGAGAGAGTGGATGCTGAA 1380
DB 1668 CTTATTTCTGACTCCCACTACATCATCAACCAAGGAGAGAGTGGATGCTGAA 1727
QY 1381 AACATCCGCTGGTAACCACTGCTGCTGGGCACTTCCACCGGCACTTCAATTCCTAC 1440
DB 1728 AACATCCGCTGGTAACCACTGCTGCTGGGCACTTCCACCGGCACTTCAATTCCTAC 1787
QY 1441 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGAGAGATCGTGGGGGCACTATCATTT 1500
DB 1788 ACCAATGAGTGGCTCCAAATAGACCTGGGGAGAGATCGTGGGGGCACTATCATTT 1847
QY 1501 CAGGTTGGGAGCAGGAGAGAGAGAGTGTTCATGAGGAGTTCAGATCGGTTACAGC 1560
DB 1848 CAGGTTGGGAGCAGGAGAGAGAGTGTTCATGAGGAGTTCAGATCGGTTACAGC 1907
QY 1561 AACAAAGCTGGGAGTGGAAAGATGATGATGAGTACAGCAACCAAGGAGGAGTCTTTT 1620
DB 1908 AACAAAGCTGGGAGTGGAAAGATGATGATGAGTACAGCAACCAAGGAGGAGTCTTTT 1967
QY 1621 GAGGCAACAACTATGATACCTGAGCTGGGAGTTCCTGAGCTCTCTCCAGCGGA 1680
DB 1968 GAGGCAACAACTATGATACCTGAGCTGGGAGTTCCTGAGCTCTCTCCAGCGG 2027
QY 1681 TTCATCAGATTTACCCCGAGAGAGCCACTCATGCGGAGTGGGGCTCAGAAATGAGCTG 1740
DB 2028 TTCATCAGATTTACCCCGAGAGAGCCACTCATGCGGAGTGGGGCTCAGAAATGAGCTG 2087

```
QY 1741 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACCGACACCTCCCAACGGGAACCTTGGTG 1800
Db 2088 CTGGGCTGTGAAGTGAAGCCCTACAGCTGGACCGACACCTCCCAATGGGAACCCAGTG 2147
QY 1801 GATGAATGTGATGACGACGAGGCACTGCCAGTGGGAACAGGTGATGATCTCCAGCTC 1860
Db 2148 CATGAGTGTGACGACGACGAGGCACTGCCAGTGGGAACAGGTGATGATCTCCAGCTC 2207
QY 1861 ACAGTGGGCACTGTGCTGGCCACAGAAAGCCACGCTCATAGACGACCAATACAA 1920
Db 2208 ACAGGAGGCACTGTCTGGCCACAGAGAAAGCCACCACTATAGACGACCAATCCAA 2267
QY 1921 TCAGAGTTTCCAAACATATGTTTAACTGTGAATTTGGCTGGGGCTCTCACAGAGCTTC 1980
Db 2268 TCAGAGTTTCCGACATACGTTTAACTGCGAGTTTGGCTGGGGCTCTCACAGAGATTC 2327
QY 1981 TGCCACTGGGAACATCACATACGTCGACGCTCAAGTGGAGTGTGTTGACCAAGAGAG 2040
Db 2328 TGCCACTGGGAGCATGACAGCCATGACAGCTCAGGTGGAGTGTGCTGACCAAGAGACA 2387
QY 2041 GGACCCATTCAGGATCACACAGAGATGGCACTTCATCTATTTCCAAAGCTGACGAAAT 2100
Db 2388 GGCCGATTCAGGACCATACAGAGATGGCACTTCATCTATTTCCAAAGCTGATGAAT 2447
QY 2101 CAGAAGGCAAGTGCGCTCGCTGGTGGAGCCCTGTGGTTTATTTCCAGAACTCTGCCAC 2160
Db 2448 CAGAAGGCAAGTAGTCCGCGCTGGTGGAGCCCTGTGGTCTATTTCCAGAGCTCTGCCAC 2507
QY 2161 TGCATGACCTTCTGGTATCACATGCTGTGGTCCAGCTCGGACACATCAGGTCACATG 2220
Db 2508 TGTATGACCTTCTGGTATCACATGCTCGGCTCTCATGTGGGTACACTGAGGTCACAACTA 2567
QY 2221 CGCTACAGAAGCCAGGAGTACGATGAGTGGTCTGGATGGCCATTTGGACACCAAGGT 2280
Db 2568 CGCTACAGAAGCCAGGAGTATGATCACTGGTCTGGATGGTGGTGGGACCAAGGA 2627
QY 2281 GACCACTGGAAGAGGCGGTGTCTGTGCTCCAAAGTCTCTGAACTTTTATCAGGTGAT 2340
Db 2628 GACCACTGGAAGAGGAGCGTGTCTGTGCTGCAAAATCTCTGAAACTATATCAGGTTAT 2687
QY 2341 TTCGAGGCGGAATCGAAAGGAACCTTGTGGATGGTGTGGATGATGATGATGAT 2400
Db 2688 TTTGAAGTGAATTCGAAAGGAACCTTGTGGATGGTGTGGATGATGATGATGAT 2747
QY 2401 AATAACCACTTCAAGAAGATTTGCAAAACCCAGCAGACCTGGATAAAAAGAACCA 2460
Db 2748 AACAACCAATTTCTCAGGAAGACTGTGCAAAACCCAGCAGACCTAGATAAAAGAACACA 2807
QY 2461 GAAATTAATTTGATGAACAGGGAGCAGCCAGGATGCAAGGTGAAGGAAGGTGAC 2520
Db 2808 GAAATTAATTTGATGAACAGGGAGCAGCTCCAGGATATGAAGGAAGGTGAC 2867
QY 2521 AAGAACATCTCCAGGAAGCCAGCAATGTGTTGAACCTTAGAACCTTAGAACCTCTCATCACC 2580
Db 2868 AAGAACATCTCCAGGAAGCCAGCAATGTGTTGAACCTTAGAACCTTAGAACCTCTCATCACC 2927
QY 2581 ATCATAGCATGAGGCGCTCGGGGTCTCTCTCGGGGGTGTCTGTGGGGTCTGTCTGTGAC 2640
Db 2928 ATCATAGCATGAGTGCCTCGGGAGTACTCTCTGGGTGACGCTGTGGAGTGTGTCTGTGAC 2987
QY 2641 TGTGCTGTGGCAATTAAGGATGTCAGAAAGAACTTGTCTGCGCCCTGGAGAACTATAC 2700
Db 2988 TGTGCTGTGGCAATTAAGGATGTCAGAAAGAACTATCTGCGCCCTGGAGAACTATAC 3047
QY 2701 TTTGAACCTTGTGGTGTGAAGTTGAAAGAAAGCAACTGAATACACAGAGTACTAT 2760
Db 3048 TTTGAACCTTGTGGTGTGAAGTTGAAAGAAAGCAACTGAATACACAGAGTACTAT 3107
QY 2761 TCGGAGGCGATGA 2772
Db 3108 TCAGAGGCGTGA 3119
```

```
RESULT 4
US-08-936-135-3
; Sequence 3, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,135
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC97-288-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2766 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-936-135-3

Query Match 79.2%; Score 2196; DB 3; Length 2766;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 2426; Conservative 0; Mismatches 340; Indels 6; Gaps 2;

QY 1 ATGGAGAGGGGGCTCGCGCTCCTCTGCGCGCTGCTCGCCCTGCTCGCCCGCGCGC 60
Db 1 ATGGAGAGGGGGCTCGCGCTGCTGCTGCGCGCCACGCTCGCCCTGCGCCCTG---GGG 57
QY 61 GCTTTTCGCAACGATGAATGTGCGGATACTATAAAATTTGAAAGCCCGGGTACCTTACA 120
Db 58 GCTTTTCGCGAGATAAATGTGCGGGACTATAAAATTTGAAACCCCGGGTACCTTACA 117
QY 121 TCTCTGTTTATCTTCTTATCACCACCAAGTGAATAATGGAATGGCTGATTCAGGCT 180
Db 118 TCTCCGGCTACCTCTTCTTACCATCCAAGTGAGAAATGGAATGGCTAATCCAAGCT 177
QY 181 CCGGACCCATACACAGAAATATGATCACTCAACCCCTCACTTCGATTTGGAGGACAGA 240
Db 178 CCGGAGCCCTACAGAGAAATATGATCACTCAACCCCTCACTTCGATTTGGAGGACAGA 237
QY 241 GACTCAAGTATGACTAGCTGGAAGTCTTCGATGGAGAAATGAAATGGACATTTAGG 300
Db 238 GACTCCAGTATGACTATGTGGAAGTATGATGAGAGAAATGAAAGTGGCCCTGTGG 297
QY 301 GGAAGTTCTGTGGAAGATAGCCCTCTCTCTGTTGTCTTTCAGGGCCATTTCTTTT 360
Db 298 GGAAGTTCTGTGGAAGATCGACCTTCACTGTGGTGTCTTTCAGGGCCATTTCTTTC 357
QY 361 ATCAAAATTTGCTCTGACTAGCAACACATGTCGAGGATTTTCCATACGTTATCAAAAT 420
Db 358 ATCAAAATTTGCTCTGACTAGCAACCCAGGGGCGAGGATTTTCCATCGGCTATGAAATC 417
```

QY 421 TTCAAGAGAGGTCTCTGAATTTCCAGAACTACAAACACCTAGTGGAGTGATAAGTCC 480
|||||
Db 418 TTCAAGAGAGGCGCGAATTTCTCAGAACTATACAGACCTACTGGAGTGATAAGTCC 477
|||||
QY 481 CCCGATTTCCCTGAAAAATATCCCAACAGCCTTGAATGCACTTATATGCTCTTTGGGCA 540
|||||
Db 478 CTTGGGTTCCCTGAAAAATACCCCAACAGCTTGGAGTGCACCTATCATCTCTTGCACCA 537
|||||
QY 541 AAGATGTCTCAGAGATTATCTCTGAAATTTGAAAGCTTTTGACCTGGAGCCTGACTCAATCT 600
|||||
Db 538 AAGATGTCTGAGATTAATCTTAGAGTTTGAAGTTTGTACCTGGAGCAAGACTCAATCTCT 597
|||||
QY 601 CCAGGGGGATGTTCTCTGCTACAGCGGCTAGAAATCTTGGGATGGATTTCCCTGATGTT 660
|||||
Db 598 CCCGAGGAATGTTCTGCTGCTATCACCCTGGAGATCTGGGATGGATTTCCCTGAGTT 657
|||||
QY 661 GGCCTTCACATTTGGCGTTACTGTGGACAGAAAACACACAGAGTTCGAATCCGATCTCATCG 720
|||||
Db 658 GGCCTTCACATTTGGCGTTACTGTGGGAGAAAACCTCTGGCCGATCCGCTCTCTTCA 717
|||||
QY 721 GGCATTTCTCTCATGTTTATTTTACACGACAGCGGATAGCAAAAGAGTTTCTTCAGCA 780
|||||
Db 718 GGCATTTCTATCCATGCTTCTTACACTGACAGCGCAATAGCAAGAGAGTTTCTTCAGCC 777
|||||
QY 781 AACTACAGTGTCTTCAGAGCAGTCTCTCAGAAAGATTTCAATGTATGGAAGCTCTGGGC 840
|||||
Db 778 AACTACAGCGTCTGCAGAGCAGCATCTCTGAAGATTTCAAGTGTATGGAGCTCTGGGC 837
|||||
QY 841 ATGGAATCAGGAGAAATTCATCTCACCAGATCAGAGTTCCTCCAGATATAGCACCAAC 900
|||||
Db 838 ATGGAATCTGGAGATCCATTTCTGACCAGATCACTGTGACTTCCAGTATGGTACCAAC 897
|||||
QY 901 TGGTCTGCAGAGCGCTCCCGCTGAACCTACCCTGAGATGGGTGGACTCCCGGAGAGAT 960
|||||
Db 898 TGGTCTGTGAGCGCTCCCGCTGAACCTACCCTGAAACGGGTGGACACGAGAGGAC 957
|||||
QY 961 TCCTACCGAGAGTGTATCAGGTAGACTTGGGCCCTTCTCGCGCTTGTGTCAGGCTCTCGGG 1020
|||||
Db 958 TCCTACCGAGAGTGTATCAGGTAGACTTGGGCCCTTCTCGCGATTCGTTACTGCTGGGG 1017
|||||
QY 1021 ACACAGGCGGCATTTCAAGAAACCAAGAAATATATATGTCAGACTTACAAAGATC 1080
|||||
Db 1018 ACACAGGTTGCCATTTCAAGAAACCAAGAAATATATATGTCAGAACTTTACAGAGTA 1077
|||||
QY 1081 GACGTTAGCTCCACGGGAAGACTGGATCAACATAAAGAGAAACAAACCTCTCTC 1140
|||||
Db 1078 GACATCAGCTCCACGGAGAGGACTGGATCACCTGAAAGGGGAATAAAGCCATTTATC 1137
|||||
QY 1141 TTTTACGGGAACACAAACCCCAAGATGTTGTTGAGTATTTCCCAAAACCACTGATA 1200
|||||
Db 1138 TTTTACGGGAACACAAATCCACAGGATGTTGTTTGGAGTTTCCCAAAACCACTGATA 1197
|||||
QY 1201 ACTCGATTTGCCGAATCAAGCTCGAATCTGGGAACTGGCATATCTATGAGATTTGAA 1260
|||||
Db 1198 ACTCGATTTGCCGAATCAAACTCGAATCTGGGAACTGGCAATATCTATGAGATTTGAA 1257
|||||
QY 1261 GTATACGGTTGCAAGATAACAGATTATCTCTGCTCTGGAATGTTGGGTATGGTGTCTGGA 1320
|||||
Db 1258 GTTATGCTGCAAGATAACAGATTACCTCTGCTCTGGAATGTTGGGATGGTGTCTGGA 1317
|||||
QY 1321 CTTATTTCTGACTCCAGATCACATCATCCAAACCAAGGAGACAGAACTGGATGCTCTGAA 1380
|||||
Db 1318 CTTATTTCTGACTCCAGATACAGCATCCAAACCAAGGAGACAGAACTGGATGCCAGNA 1377
|||||
QY 1381 ACATCCGCTGGTAACCACTGCTGCTGGTGGGCACTTCACCCGCACTCATTTCTCTAC 1440
|||||
Db 1378 AACATCCGCTGGTGAACCACTGCTGCTGGGCACTTCACCCGCACTCATTTCTCTAC 1437
|||||
QY 1441 ATCAATGAGTGCCTCAATAGACCTGGGGAGGAGAGATGCTGAGGGGCATCATATT 1500
|||||
Db 1438 ATCAATGAGTGCCTCAATGGAOCTGGGAGATGAGAGATAGTAAGAGGTGTCTCATATT 1497
|||||

QY 1501 CAGGTGGGAAGACCCGAGAGAAACAGGTGTTTCATGAGGAAGTTTCAAGATCGGTACAGC 1560
|||||
Db 1498 CAAGGTGGGAAGACCCGAGAGAAACAAAGTGTTCATGAGGAAGTTTCAAGATCGCCTACAGT 1557
|||||
QY 1561 AACAAGCGTCTGGACTGGAAGATGATCATGATGATACAGCAACGCAAGCGGAAGTCTTTT 1620
|||||
Db 1558 AACAATGTTCTGACTGGAAATGATCATGATGATGATGATGATGATGATGATGATGATGAT 1617
|||||
QY 1621 GAGGGCAACAACAATATGATACCTGAGCTGCGGACTTTTCCAGCTCTCTCCACGGCA 1680
|||||
Db 1618 GAAGCAACAACAATATGACACACTGAGCTCCGGGCTTTTACACCTCTCTCCACAAGA 1677
|||||
QY 1681 TTCATCAGGATCTACCCGAGAGAGCCACTCATGCGGAGCTGGGCTCAGAAATGGAGCTG 1740
|||||
Db 1678 TTCATCAGGATCTACCCGAGAGAGCCACACATAGTGGCTCGGACTGAGGATGGAGCTA 1737
|||||
QY 1741 CTGGGCTGTGAAGTGGAGCCCTTACAGTGGACCGACACTCCCAAGCGGGAACCTTGGTG 1800
|||||
Db 1738 CTGGGCTGTGAAGTGAAGTGCCTACAGTGGACCGACACCACTTGGGAACCCCGTG 1797
|||||
QY 1801 GATGAATGTGATGACGACAGCCAGCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
|||||
Db 1798 GACGAGTGTGACGATGACAGGCCAACTGCCACAGTGGCACAGGTGATGACTTCCAGCTC 1857
|||||
QY 1861 ACAGTGGCACACACTGTCTGGCCACAGAAAAGCCCGTCAATAGACAGCACCATACAA 1920
|||||
Db 1858 ACAGGAGGCACCACTGTCTTGGCCACAGAGAAGCCCACTTATAGACAGCACCATCCA 1917
|||||
QY 1921 TCAGAGTTTCCACATATGTTTAACTGTGAATTTGCTGGGCTCTCCACAGACCTTC 1980
|||||
Db 1918 TCAGAGTTTCCGACATACGGTTTAACTGCGAGTTTGGCTGGGCTCTCCACAGACATTC 1977
|||||
QY 1981 TGCCACTGGGAACATGACAATCAGCTGCAGCTCAAGTGGAGTGTGTTGACCAAGCAAGAG 2040
|||||
Db 1978 TGCCACTGGGAACATGACAGCCACCGCAGCTCAGGTGGAGGTGCTGNACCAAGAGAG 2037
|||||
QY 2041 GGACCCATTCAGGATCAGACAGGAGATGGCAACTTCATCTATTCCCAAGCTGACAAAAT 2100
|||||
Db 2038 GGGCCCATTCAGGACACACAGGAGATGGCAACTTCATCTATTCCCAAGCTGATGAAAAT 2097
|||||
QY 2101 CAGAAGGCAAGTGGCTGGCTGGTGAGCCCTGTTGTTTATTTCCAGACACTCTGCCAC 2160
|||||
Db 2098 CAGAAGGCAAGTGGTGAGCCCTGTTGAGCCCTGTTGTTTATTTCCAGAGTTCTGCCAC 2157
|||||
QY 2161 TGCATGACTTCTGCTATCACATGCTGGTCCACGCTCGGCACACTCAGGTGCAAACTG 2220
|||||
Db 2158 TGCATGACTTCTGCTATCATGCTCGGCTCTCATGTTGGGTACACTGAGGTGCAAACTG 2217
|||||
QY 2221 CGCTACCAAGCCAGAGAGTACGATCAGCTGGTCTGGATGGCCATTTGGACACCAAGGT 2280
|||||
Db 2218 CACTACCAAGCCAGAGAGATATGATCAACTGGTCTGGATGGTGGTGGCCACCAAGGA 2277
|||||
QY 2281 GACCACTGGAAGGAGGGGTGCTTGTCTCCACAAGTCTCTGAAAATTTATCAGGTGATT 2340
|||||
Db 2278 GACCACTGGAAGGAGGGGTGCTTGTCTGTCACAAATCTCTGAAAATCTGATCAGGTATT 2337
|||||
QY 2341 TTCAGGCGGAATCGGAAAAAGAAACCTTGGTGGATTGCTGTGATGATGATGATGATGAT 2400
|||||
Db 2338 TTTGAAGTGAATTCGGAAGAAACCTCGGTGGGATGCTGTGGATGATGATGATGATGAT 2397
|||||
QY 2401 AATAACCACTTTTCAAGAAGATTGTGCAAAACAGACAGACTGGATAAAAAGAACCCA 2460
|||||
Db 2398 AACAACCACTTTCTCAGGAGGACTGTGCAAAACCAACAGACTAGATAAAAAAGAACACA 2457
|||||
QY 2461 GAAATTTAAATTTGATGAACAGGAGGAGCGGACAGGATAGGAAGGTGAAGGAAAGGTGAC 2520
|||||
Db 2458 GAAATTTAAATTTGATGAACAGGAGGAGCGGACCGGAGGATGTA --GAAGGGAAGGCGAC 2514
|||||
QY 2521 AAGACATCTCCAGGAGCGGCAATGTTGTAAGACCTTTAGAACCTTAGAACCTCTCTCATCACC 2580
|||||
Db 2515 AAGACATCTCCAGGAGCGGCAATGTTGCTTAAGACCTGGGACCTCTCTGATCACC 2574
|||||
QY 2581 ATCATAGCCATGAGCGCCCTGGGGGTCTCTCTGGGGGTGCTGTGGGGTCTGCTGCTGAC 2640
|||||

Db 2575 ATCATAGCCATGAGTCCCTGGGGGCTCTCTGGGTGCACTCTGTGGAGTGTGCTGTAC 2634
QY 2641 TGTGCTGTGGCATATATGGGATGTCAGAAAGAACTTGTCTGCCCTGGAGAACTATAAC 2700
Db 2635 TGTGCTGTGGCACAATGGGATGTCGGAAGAACTATCTGCCCTGGAGAACTATAAC 2694
QY 2701 TTTGAACCTTGTGGATGGTGTGAAGTTGAAAAAGACAACTGAATACACAGAGTACTTAT 2760
Db 2695 TTTGAACCTTGTGGATGGTGTGAAGTTGAAAAAGATAAAGTGAACCAACACACATTAATAC 2754
QY 2761 TCGAGGCGATGA 2772
Db 2755 TCAGAGGCGGTGA 2766

RESULT 5

US-08-936-135-11
; Sequence 11, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,135
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC97-288-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-936-135-11

Query Match 21.5%; Score 596.2; DB 3; Length 4733;
Best Local Similarity 54.4%; Pred. No. 4.1e-169;
Matches 1466; Conservative 0; Mismatches 1143; Indels 84; Gaps 10;

QY 107 CCGGGTACCTTACATCTCTCGTGTATCTCTTATCCATCTTATCACCAAGTGAATAATGCGAAT 166
Db 676 CTGGCTACATCACTTCCCGAGGTACCCCGAGGACTATCCCTCCACCAAGACTGTGACT 735
QY 167 GCGTGTACGCTCGGACCCATACAGAGATTAATGATCACTTCAACCTCACTTCG 226
Db 736 GGATTTGCTACGCCCCCGAACCCACAGAGATTGTTCTCAACTTCAACCTCACTTTCG 795
QY 227 ATTTGGAGACAGAGACTCAAGATGACTACGTGGAAGTCTTCGATGGAGAAATGAAA 286
Db 796 AAATCGAAGAACAGACTCAAGATGACTTTCATTGAGATTCGGGATGGGACAGTGA 855

QY 287 ATGGACATTTTAGGGAAAGTTCTGTGAAAGATAGCCCCCTCTCTGTGTGTCTTTCAG 346
Db 856 CAGCTCAGCTCTCTGGCAAGCACTGTGGGAACATCGCCGCCACCATCATCTCTCTCAG 915
QY 347 GGCATTTCTTTTATCAAAATTTGTCTGTACTACGAAACACATGTTGCGAGGATTTTCCA 406
Db 916 GCTCCGTGTATACATCAAGTTCCACTCAGACTACGCCCGGAGGGGGAGGTTTCTCTC 975
QY 407 TAGCTTATGAAATTTTCAAGAGAGTCTCTGAA-- --TGTTCACCAAGACTACACACACCTA 463
Db 976 TAGCTATGAGATCTTCAAAACAGGCTCTGAAGATTGTTTCCAAAGAACTTTACAAAGCCCCA 1035
QY 464 GTGGAGTGAATAAGTCCCGCGGATTCCTTGAAATAATCCCAACAGCTTGAATGCACTT 523
Db 1036 ATGGGACCATTTGAATCTCCAGGGTTTCCAGAGAAGTATCCACACAATCTGGAGTGTACCT 1095
QY 524 ATATTCCTTTTGGCCAAAGATGTACAGATTAATCTCGAATTTGAAAGCTTTGACCTGG 583
Db 1096 TCACCATCTCTGGCCAAACCCAGGATGGAGATCATCTACAGTTTCTCTGACCTTTGACCTGG 1155
QY 584 AGCCTGACTCAAAATCTCCAGGGGGATGTTCTGTCTGCTACGACCGGCTAGAAATCTGGG 643
Db 1156 AGCATGACCCCTTACAAAGTGGGGGAAGAGACTGTAAATATGACTGGCTGGAGATCTGGG 1215
QY 644 ATGGATTCCCTGATTTGGCCCTCACAATTTGGGCGTTACTGTGGACAGAAACACCAAGTTC 703
Db 1216 ATGGCATTCACATGTTGGACCTCTGATTTGGCAAGTACTGTGGACGAAACACCTCCCA 1275
QY 704 GAATCCGATCTCATCGGGCATTTCTCCATGTTTATACCCGACGAGCGGATAGCAA 763
Db 1276 AACTCGCTCGTCCACGGGGATCTCTCTCTGACCTTTTACACGGGACATGGCAGTGGCCA 1335
QY 764 AAGAAGTTTCTCAGCAAACTACAGTGTCTGACAGAGCAGTGTCTCAGAAGATTTCAAT 823
Db 1336 AGGATGGCTTCTCCGCACTTACTATTGATCCACAGGAGCACCCTGAGATTTTCAGT 1395
QY 824 GTATGGAAGCTCTGGCCTGGAATCAGAGAAATTCATTCTGACAGACTACACAGCTTCTT 883
Db 1396 GCAATGCTCTTTGGGAATGGAGTCTGGCCGGATTCTTAATGAACAGATCAGTGCCTCT 1455
QY 884 CCAAGTA---TAGCACCAACTGTCTGACAGAGCGCTCCCGCTGAACTACCTTGGAATG 940
Db 1456 CCACCTTCTCTGATGGGAGTGGACTCTCAACAGAGCGCGCTCCATGTTGATGACAATG 1515
QY 941 GTGGACTCCCGAGAGGATTTCTACCGAGAGTGGATACAGGTAGACTTTGGGCTTCTGC 1000
Db 1516 GCTGGACACCAATTTGGATTCCAAAGAGATATCTCCAGGTGGACCTCGCTTCTCTAA 1575
QY 1001 GCTTTGTACGGCTGTGGGACACAGGGCGCCATTTCAAAGAAACCAAGAAATATTT 1060
Db 1576 CCATGCTCACAGCCATTGCAACACAGGGAGCCATTTCAGGGAAACCCAGAAAGGCTACT 1635
QY 1061 ATGTCGAAGACTTACAAGATCGAGTTAGTCTCAACGGGGAAGACTGGATCACCATAAAG 1120
Db 1636 ACCTCAAAATCGTACAAGTGGAGTCAACACAAATGTTGAAGATTGGATGGTCAACGGC 1695
QY 1121 AAGGAAACAAACCTGTCTCTTTTCAGGGAACACCAACCCACAGATGTTGTGGTTGCAG 1180
Db 1696 ATGGCAAAACCAAGATATTTCCAAAGCGAACATGATCGGAGGAGTGGTGTAAACA 1755
QY 1181 TATTTCCCAAAACCTGATACTCGATTTGTCCGAATTAAGCTTGAAGTGTGGAACTTGG 1240
Db 1756 AGCTCCACATGCACTGTGACTCGGTTCATCAGGATCGCCCGCAGAGCTGSCATTTGG 1815
QY 1241 GCATATCTATGAGATTTGAAGTATACGGTTTGAAGATACAGATTAATCTTCTCTGGA 1300
Db 1816 GCAATGCTTCCCTCGGCTGGAGTCTTTGGTCCGGGTACAGATGACATGCTGCTCAACA 1875
QY 1301 TGTGGGATGTTGTCTGGACTTATTTCTGACTTCCAGATACATCATCATCAACCAAGGAG 1360
Db 1876 TCTTGGGATGCTCTCGGGCTCATTTGCTGATACCCAGATCTCTGCTCTCTCCACCCGAG 1935

Qy	1361	ACAGAAAC	TGGATG	CCTG	AAAA	CATCCG	CGTGGT	GTAA	CCAG	TGCT	CGCT	TGCG	TGGCG	CACTTC	1420
Db	1936	AGTACCT	CTGG	AGCC	CCAG	TGCTG	CCG	CGCT	GGT	TAGT	AG	CCGCT	TGCG	TGGTT	CCCTC
Qy	1421	CACCGC	CACTC	ATTCC	TACAT	CAAT	GAGT	GGCT	CCAA	TAGAT	AG	ACCT	TGGG	GAGGA	GA
Db	1996	GGAA	CCCTCA	AGCC	CCAG	CGTGA	AGAA	TGCTT	CAGG	TAG	AGCT	GGG	CA	CCCA	GA
Qy	1481	TCGT	GAGGG	GCAT	CATTC	ATTC	AGGT	GGG	AA	GCAC	CCG	GAGA	-----	-----	1522
Db	2056	CAGT	GAA	AGGG	GT	CATC	ATCC	AGG	GAG	CCG	GAG	GAG	CA	GAC	TAC
Qy	1523	ACAAG	GTG	TTCAT	GAG	GAAG	TTC	AAG	ATG	CGG	GT	CAG	CA	CA	CG
Db	2116	CCAG	CGG	CTT	GTG	TAC	CAAG	TTC	CAAG	TCT	CC	TAC	AG	CG	CT
Qy	1583	TGAT	CAT	GGAT	GAC	AGCA	CAAC	CAAG	CGG	CA	GTCT	TTT	TG	AGG	CA
Db	2176	ATAT	CCAG	AG	CC	AG	CACT	CAG	CAG	CA	AA	GTCT	TTT	TG	AGG
Qy	1643	CACCT	GAGCT	GGG	ACT	TTTT	TCC	AGCT	CTCT	CC	AC	CG	GA	TT	CCAT
Db	2236	CCCT	GACAT	CCG	AG	TT	CGAT	CTCT	GT	TCC	AG	CG	CA	GT	AT
Qy	1703	GAGC	CACT	CAT	TG	CG	GA	CT	GGG	CT	CAG	AT	TG	CA	AT
Db	2296	GGT	GGT	CG	CA	G	CG	AT	CGG	AT	G	AG	CT	G	CT
Qy	1763	CTAC	AGT	GG	AC	CG	CACT	TCC	AC	GG	GA	CTT	GGT	GG	AT
Db	2356	CAA	AG	CC	CA	AGT	GG	AG	CG	-----	TGG	AG	CC	AC	CG
Qy	1823	CCA	ACT	GC	CA	CA	GTG	AT	GA	CTT	CC	AG	CT	CAC	AG
Db	2412	CCAT	AT	CC	CA	TG	AT	GA	GT	AG	CA	GT	AG	CA	GT
Qy	1883	CCAC	AGAA	AG	CC	CA	CGT	CAT	AG	CA	GC	AC	CA	AT	CA
Db	2450	AA	ACT	GC	AG	CTT	GG	AT	G	CA	AA	GA	TT	G	CA
Qy	1943	TTAA	CT	GT	GA	TT	TG	CT	GG	GC	CT	CT	CA	AG	AC
Db	2502	-----	TG	CA	ACT	TT	TG	AT	TT	CC	GG	AG	AG	AC	CT
Qy	2003	AC	GT	CA	GT	CA	AGT	GT	TT	G	CA	CA	AG	CG	CA
Db	2551	CCA	AGT	GG	CT	CG	GA	CA	CGT	GG	AT	CG	CA	CG	CA
Qy	2063	GAG	AT	GG	CA	ACT	T	CT	AT	TT	CC	CA	AGT	CA	AG
Db	2611	AT	G	CA	AA	GA	ACT	TT	CT	T	G	AA	CT	G	CA
Qy	2123	TG	GT	GA	CG	CT	GT	TT	AT	TT	CC	CA	AG	AC	CT
Db	2671	TC	AT	CA	GG	CC	CA	CG	GT	GC	AC	CT	GT	GT	CA
Qy	2183	TG	CT	TG	GT	CC	CA	CT	CG	CA	CA	CT	CG	CT	CA
Db	2731	CC	AT	GG	CG	CC	AG	GG	GT	GG	CA	CT	GC	AG	GT
Qy	2243	AC	GA	CT	CA	GT	TG	CT	GG	AT	GG	CA	CT	GG	CA
Db.	2782	AA	AG	CA	ACT	TC	CT	TT	GG	GT	CA	TC	CG	T	AG
Qy	2303	TC	TT	GT	CC	CA	AG	CT	CT	GT	AA	CT	TT	AT	CA
Db	2842	TT	AT	CT	CG	CC	CA	CT	AT	CA	GT	AG	AT	CA	GT
Qy	2363	GA	AA	CT	TT	GT	GG	AT	TT	GT	GG	AT	TT	AT	AA
Db	2902	G	AG	AT	CG	GG	AG	AT	TT	CC	GG	AT	GA	CT	TT
Qy	2423	ATT	GT	GC	AA	AC	CA	GA	CG	AC	CT	GG	AT	AA	AA

[illegible]

RESULT 6

```

US-08-936-135-9
: Sequence 9, Application US/08936135
: Patent No. 6054293
: GENERAL INFORMATION:
: APPLICANT: Tessier-Lavigne, Marc
: APPLICANT: He, Zhigang
: APPLICANT: Chen, Hang
: TITLE OF INVENTION: Semaphorin Receptors
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 75 DENISE DRIVE
: CITY: HILLSBOROUGH
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94010
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/936,135
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: OSMAN, RICHARD A
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: UC97-288-2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 343-4341
: TELEFAX: (650) 343-4342
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4718 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-936-135-9

```

Query Match

Query Match 21.0%; Score 583; DB 3; Length 4718;

Query Match: 21.0%, Score 305, DB 3,
Best Local Similarity 54.78; Pred, No. 3.9e-165;

Best Local Similarity	Seq. ID	Accession No.	Length	Score	Indels	Gaps
Matches 1474; Conservative	0	Mismatches 1120;	99	12;		

CATPBA TOCINO
CATPBA TOCINO

QY 107 CCGGTACCTTACATCTCCTGGTTATCCTCATTTCTATCACCCAAAGTGAAAAATGCGAAT 166

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 676 CTGGCTACATCACTTCCCCAGGCTACCCCCAGGACTATCCCTCCACCACTGTGAGT 735

Db 2842 TTATCCTCCCGAGCTATGACATGGAGTATCAGATCGTCTTCGAGGGAGTATAGGAAGG 2901
Qy 2363 GAAACCTTTGGTGGGATGCTGTGGATGACATTAGTATTAAATAAACACACATTTACAGAAG 2422
Db 2902 GACGATCGGGAGAGATTTCCGGGATGACATTCGGATAGCACTGATGTCCCACTGGAGA 2961
Qy 2423 ATTGTGCAAAACAGCAGACCTGGATNAAAAAGAACCCAGAAATTTAAAATTGATGAACAG 2482
Db 2962 ACTGCATGAAACCCATATCAGCTTTTGAGA-----TGAAATATGAAGAGATT 3009
Qy 2483 GGAGCACCAGGATACAAAGGTGAAGAGAGGTGACAAAGACATCTCCAGGAAGCCAG 2542
Db 3010 GGAGCAACTCTCTCTCTCTACCTCAGGGGCTGGTGAC---CCCTCATCTGGCAAGAAA 3066
Qy 2543 GCAATGTGTGAAGACCTTAGAACCCATCCTCATCACCATCATAGCCATGAGCGCCCTGG 2602
Db 3067 AGAGCTGCTGTACACCTAGATCCCATCTGTATCACCATCATCGCCATGAGCTCGCTGG 3126
Qy 2603 GGGTCTCTCTGGGGCTCTCTGTGGG---TCGTGCTGTACTGTGCTGTGGCATAATG 2659
Db 3127 GGGTCTCTCTGGGGCCACCTGTGGGGCCCTCTCTCTTTACTGCACTGCTCCTATTCTGG 3186
Qy 2660 GATGTGCAAGAAACATTTGCTGCCCTGGAGAACTATACTTTGAACTTTGGATGGTG 2719
Db 3187 GTCTGAGTTCGAGGAGCTGCACCACTGGAGAACTACAACTTTGAGCTCTACGATGCC 3246
Qy 2720 TGAAGTTGAAAAAGACAACTGAATACACAGAGTACTTATTCGGAGGCATGA 2772
Db 3247 TCAA---GCACAGGTCAAGATCAATCATCAGAAAGTGTCTCGGAGGCATGA 3296

RESULT 7

US-08-936-135-7
; Sequence 7, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936.135
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC97-288-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-936-135-7

Query Match 21.0%; Score 581.4; DB 3; Length 3539;
Best Local Similarity 54.7%; Pred. No. 1e-164;
Matches 1473; Conservative 0; Mismatches 1121; Indels 99; Gaps 12;

Qy 107 CCGGTACTCTTACATCTCCTGGTATCTCTTATCTCTTATCACCACCAAGTGAATAATCGGAAT 166
Db 677 CTGGTACATCACTTCCCGAGGTACCCCGAGGACTATCCCTCCACCCAGAACTGTGAGT 736
Qy 167 GCGTGTACAGCTCCGGACCCATACAGAGAAATATGATCAACTTCAACCTCACTCTCG 226
Db 737 GATTGTCTAGCCGCCCAACCAACAGAAATTTCTCAACTTCAACCTCACTCTT 796
Qy 227 ATTTGGAGACAGAGCTGCAAGTATGACTAGCTGGAAGTCTTCTGATGAGAAAATGAAA 286
Db 797 AAATCGAAGAACACGACTGCAAGTATGACTTCATTTGAGATTCGGGATGGGACAGT 856
Qy 287 ATGGACATTTTATGAGGAAAGTTCTGTGGAAGATAGCCCCCTCTCTCTCTGTGTCTT 346
Db 857 CAGCTGACCTCTCTGGGCAAGCAGCTGTGGGAACTATCGCCCGCCCACTCATCTCTCT 916
Qy 347 GGCATTTCTTTTATCAAAATTTGCTCTGACTAGCAAGAACACATGGTCAGAGTTTCCA 406
Db 917 GCTCGTGTATACATCAAGTTTCACTCAGACTACGCCCGGAGGGGCGAGGTTTCTCT 976
Qy 407 TACGTTATGAAATTTTCAAGAGAGGTCTCTGAA---TGTTCAGAGAACTACACACACCTA 463
Db 977 TACGCTATGAGATCTTCAAAACAGGCTCTGAAGATTGTTCCAGAACTTTACAGCCCA 1036
Qy 464 GTGGAGTGAATAAGTCCCGGATTCCTGTGAAAAATATCCCAACAGCCTTGAATGCACT 523
Db 1037 ATGGAGCAATTTGAATCTCCAGGGTTTCCAGAGAACTTCCACAACTTGGACTGTACT 1096
Qy 524 ATATTGTCTTTGCGCCAAAGATGTGAGAGATTATCTCGAATTTCAAGACTTTGACCTGG 583
Db 1097 TCACCATCTGGCCAAACCCAGGATGGAGATCATCTACAGTTCTCTGACCTTTGACCTGG 1156
Qy 584 AGCCTGACTCAAACTCTCCAGGGGGATGTCTGCTGCTAGCAGCGGTAGAAAATCTGGG 643
Db 1157 AGCATGACCTCTACAGTGGGGAGGAGACTGTAAATATGACTGGCTGGACATCTGGG 1216
Qy 644 ATGGATCCCTGATGTGGCCCTCACATTTGGCGTACTGTGGACAGAAAACACACAGTCT 703
Db 1217 ATGGCATTTCCACATTTTGGACCTCTGATTTGGCAAGTACTGTGGGACAGAAAACACCTCCA 1276
Qy 704 GAATCCGATCTCATCGGGCATTTCTCTCACTGGTTTTTACACCGACAGCGGATAGAA 763
Db 1277 AACTCCGCTCTGCCACGGGGATCTCTCTTGTACCTTTTACACCGGACATGGCAGTGGCCA 1336
Qy 764 AAGAAGTCTCTCAGCAAACTACAGTGTCTGCGAGAGCAGTGTCTCAGAAATTTCAAA 823
Db 1337 AGGATGGCTTCTCCGACGTTACTATTGATCCACGAGGACCCACCTGAGAAATTTTCA 1396
Qy 824 GTATGGAAGCTCTGGGCATGGAATCAGAGAAATTTCTGACACAGATCAGAGTCTTT 883
Db 1397 GCAATGTCCCTTTGGGAATGGAGTCTGGCCGATTTGCTTAATGAAACAGATCAGTGGCTCT 1456
Qy 884 CCCAGTA---TAGCACCAACTGGTCTGAGAGCGCTCCCGCTGAACTACCTGAGAAATG 940
Db 1457 CCACCTTCTGTGATGGGAGGTGGACTCTCTCAACAGAGCGGCTCCATGGTATGACAATG 1516
Qy 941 GGTGGACTCCCGGAGAGGATTTCTTACCAGAGATGGATACAGATAGACTTTGGGCCCTTCTGC 1000
Db 1517 GCTGGACACCCAAATTTGGATTCCCAACAGGAGTATCTCCAGGTGGACCTGGCTTCTCTAA 1576
Qy 1001 GCTTTGTACGGCTGTGGGACACAGAGCGGCTTTCAAAAAGAAACCAAGAAAGAAATAT 1060
Db 1577 CCATGCTCACAGCCATTGCAACACAGAGGAGCCATTTCCAGGGGAAACCCAGAAAGGCTACT 1636
Qy 1061 ATGTCAAGACTTACAGATCGACTTAGCTTCAACGGGGAGAGCTGGATCACCATAAAG 1120
Db 1637 ACGTCAAAATCGTACAAGCTGGAAGTTCAGACAAAATGGTGAAGTTGGATGTTCTACCGGC 1696


```
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2730 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-936-135-17

Query Match      20.6%; Score 570.6; DB 3; Length 2730;
Best Local Similarity 54.5%; Pred No.1.5e-161;
Matches 1483; Conservative 0; Mismatches 1139; Indels 99; Gaps 13;

QY 79 TGTGGCGATPACTATAAAAATTGAAAGCCCGGGTACCTTTACATCTCTGTTATCCTCAT 138
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 82 TCGGGAGGTCGTTGAATTCGAAGATGCTGGCTATATCACTCTCCCGGTACCCCGAG 141
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 139 TCTTATCACCCCAAGTGAANAATGCGAATGGCTGATTCAGGCTCCGGACCCATACAGAGA 198
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 142 GACTACCCCTCCACAGAAACTGCGAGTGGATTGTTTACGCCCGGAAACCCAAACAGAA 201
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 199 ATTATGATCAACTTCAACCCCTCACTTCGATTTGGAGGACAGACTGCAAGTATGACTAC 258
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 202 ATTGCTCTCACTTCAACCCCTCACTTGAATCGAGACAGACTGCAAGTATGACTTT 261
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 259 GTGGAAGTCTTCGATGGAGAAAATGAAAATGGACATTTTATGGGGAAGTTCTGTGGAAG 318
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 262 ATCGAGATTCGGGATGGGACAGTGAATCCGCAGACCTCTCTGGGCAACACACTGTGGAA 321
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 319 ATAGCCCTCTCTCTGTGTCTTCAGGGCCATTTCTTTTATCAAAATTTGTCTCTGAC 378
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 322 ATGCGCCCGCCACCATTATCTCTCGGGCTCCATGCTCTACATCAAGTTCACTCCGAC 381
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 379 TACGAAACACATGGTGCAGGATTTTCCATACGTTATGAAATTTTCAAGAGAGTCTCTGAA 438
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 382 TAGCCCGCGAGGGGAGGCTTCTCTGCGCTAGGAGATCTTCAAGACAGCTCTGAA 441
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 439 ---TGTTCGCAAGCTTACACACACCTAGTGAGTGATGAAAGTCCCGCGGATTCCTGAA 495
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 442 GATTGCTCAAAAACCTTCAAGGCCCAACGGGACCATCGAATCTCTCGGTTTCTCTGAG 501
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 AAATATCCCAACAGCCTTGAATGCACCTATATGTCTTTGCGCCCAAGATGTCAGAGATT 555
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 502 AAGTATCCACAACTTGGAGCTGACCTTTACCATCTCTGGCCCAACCCCAAGATGGAGATC 561
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 556 ATCCTGGAATTTGAAAGCTTTGACCTGGAGCTGACTCAAAATCTCCAGGGGGGATGTC 615
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 562 ATCCTGCAGTCTCTGATCTTTGACCTGGAGCATGACCTTTGCAGGTGGGAGGGGGAC 621
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 616 TGTGCGTAGCAGCCGCTAGAAATCTGGGATGGATTCCCTGTATGTTGGCCCTCACATTTGG 675
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 622 TGCAAGTAGGATGGCTGGACATCTGGGATGGCATTCACATGTTGGCCCCCTGATTGGC 681
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 676 CGTTACTGTGGACAGAAAACACAGGTGCAATCCGATCTCATCGGGGATTTCTTCCATG 735
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 682 AAGTACTGTGGGACAAAACCCCTCTGAACTTCGTTTATCGACGGGATCTCTCCCTG 741
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 736 GTTTTTTACACGACAGCCGATAGCAAAAGAGTTTCTTCAGCAAACTACAGTGTCTTG 795
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 742 ACCTTTCACACCGCATGCGGTGGCCAGGATGGCTTCTCTGCGGTACTACCTGGTC 801
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 796 CAGAGCAGTGTCTCAGAAAGATTCAAAATGATGGAAGCTCTGGGCATGGAATCAGGAGAA 855
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 802 CACCAAGAGCCACTAGAGAACCTTTCAGTGAATGTTCCCTCTGGGCATGGAGCTGSCCGG 861
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 856 ATTCACTTCGACCATGACAG---CTTCTTCCAGTATAGCACCAACTGGTCTGCAGAG 912
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 862 ATTGCTAATGAACAGATCAGTGCCTCATCTACTCTGATGGGAGGTGGACCCCTCAA 921
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 913 CGCTCCGCTGAACTACCTCAGAGATGGTGGACTCCCGGAGAGGATTCCTACCGAGAG 972
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 922 CAAGCCGCTCCATGGGTGATGACAAATGGCTGGACCCCAACCTTGGATTCCCAACAGGAG 981
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 973 TGGATACAGGTAGACTTGGGCCTTCTGCGCTTTGTACAGCGCTGTCCGGACACACAGGCGCC 1032
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 982 TATCTCCAGGTGGACCTGCGCTTTTAAACCATGCTCAGGCCATCGCAACACAGGAGCG 1041
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1033 ATTTCAAAAGAAACCAAGAAATATTTATGTCACAGACTTACAGATCATGAGCTAGCTCC 1092
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1042 ATTTCCAGGNAACACAGAATGGCTACTAGCTAAATCTACAAGCTGGAAGTGCACACT 1101
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1093 AACGGGAAGACTGGATACCATATAAAGAAAGAAACAAACCTGTCTCTTCAGGGAAC 1152
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1102 ATGAGAGAGACTGGATGTTGTTACCGGCATGGCAAAACCAACAGTATTTCAAGCCAAC 1161
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1153 ACCAAACCCACAGATGTTGTTGTTGCGATTTTCCCAAAACCACTGATAACTCGATTGTC 1212
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1162 AACGATGCAACTGAGGTGTTCTTGAACAAGCTCCACGCTCCACTGCTGACAAGTTGTT 1221
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1213 CGAATCAAGCCTGCAACTTGGGAAACTGCGATATCTATGAGATTTGAAATATACGTTGC 1272
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1222 AGAATCCGCGCTCAGACCTGGCACTCAGATATCGCCCTCCGGCTGGAGCTTTCGGCTGC 1281
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1273 AAGATAACAGATTATCTTCTGCTTGAATGTTGGGTATGTTGCTGCTGACTTATTTCTGAC 1332
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1282 CCGGTACAGATGCTCCCTGCTCCAACATGCTGGGATGCTCTCAGGCTCATTCGACAG 1341
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1333 TCCCAGATCACATCATCCAACCAAGAGACAGAAACTGGATGCTGTGAAACATCCGCTG 1392
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1342 TCCAGATCTCCGCTCTTCCACCCAGGAATACCTCTGGAGCCCACTGCAGCCGCCCTG 1401
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1393 GTAACAGTCTGCTGGCTGGGACATTCACCGGCACCTTCATTCATCATCAGGAGCCCG 1452
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1402 GTCAGCAGCGCTCGGGCTGTTTCCCTCGAATCCCTCAGGCCCCAGCCCGGTGAGGAGT 1461
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1453 CTCCTCAATAGACTGGGGGAGGAGAACTGCTGAGGGGATCATCATTCAGGCTGGGAAG 1512
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1462 CTTCAGGTAGATCTGGGAACACCAACAGACAGTGAAGGTGTCATCATCAGGAGCCCG 1521
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1513 CACCGAGAGAA-----CAAGGTGTTTCATGAGGAAGTTCAAGATCGGG 1554
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1522 GGAGGAGACAGATACACTGCTGTGGAAGCAGAGCATTTTGTGCGCAAGTTCAAAGTCTCG 1581
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1555 TACAGCAACACCGCTCGGACTTGAAGATGATCATGTGATGACAGCAACAGCGAGGGAAG 1614
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1582 TACAGCCTTAACGGCAAGCACTGGGAATACATTCAGGACCCCAAGGACCCAGCAAG 1641
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1615 TCTTTTGGGGCAACAACTATGATACACCTGAGCTGCGGACTTTTCAGCTCTCTCC 1674
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1642 CTGTTGGAAGGAACTATGACATGACACCTGACATCGAAGGTTTGACCCCATTCG 1701
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1675 AGCGATTCATCAGGATCTACCCCGAGAGAGCCACTCATGCGGAGCTGGGGCTCAGAA 1734
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1702 GCACAGTATGTGCGGGTATACCCGGAGAGGTGGTCCCGCGGGGATGGGATCGGCTG 1761
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1735 GAGCTGCTGGGCTGCAAGTGGNAGCCCTACAGCTGGAGCCGACCTCCCAACGGGAAC 1794
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1762 GAGGTGCTGGGCTGTGACTGGACAGACTCCAAGCCCAAGTAAAAACGC-----TGGGAC 1817
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1795 TTGGTGGATGAATGATGACGACCAAGGCAACTGCGCAGGTGGAACAGGTGATGACTTC 1854
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1818 CACTGGAAGAGCGGAAGAGACAAACCCCTACCCCAAGAGAGAGGCCA----- 1870
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1855 CAGCTCAGAGGTGGCACTGCTGTGGCCACAGAAAGCCCAAGCTCATAGACAGACC 1914
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1871 CAGAGTGTGGGGAAGTGCAGCTTTGAGGATGACAAG-----ATTTGCAAGCTCG 1921
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1915 ATACAATCAGATTTTCCAACATATGTTTAACTGTGAATTTGGCTGGGCTCTCACAAG 1974
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1922 CTTGGGATTCATTCGAA-----CTTCGATTTCTTCGAG 1956
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1975 ACCTTCTGCCACTGGGAACATGACATCAGTGCAGCTCAAGTGGAGTGTGTTGACGAG 2034
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1957 GAGCCCTGTGGTGGATGTATGACCATGCCAAGTGGCTCCGAGACCACTGGGGCAGCAG 2016
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2035 AAGACGGGACCCATTTCAGGATCACACAGAGATGGCAACTTCATCTATTTCCAAGCTGAC 2094
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 2017 TCAGGCCAAACACCGGACGTTTCAGATGACAGAAATTTCTTGGCGTGCAGAGTGAC 2076
QY 2095 GAAATACAGAAAGTGGCTCCCTGGTGAGCCCTGTGTTTATTCACCAACTCT 2154
Db 2077 AGCCAGAGAGGCGCAGTATGCCCGGCTCATCAGCCCGCTGTCCACCTGCCCGCAAGC 2136
QY 2155 GCCACTGTCATGACCTTCTGGTATCATCTGTGGTCCACATCGCGCACACTCAGGGTC 2214
Db 2137 CCGGTGTGATGAGTTCCAGTACCAGGCCACGGCGCGCGGGTGGCGCTGCAGGT - 2195
QY 2215 AACTCGGCTACAGAACCCAGAGAGTAGCATCAGCTGTGTGATGCGCATTCAGACAC 2274
Db 2196 -----GGTGGCGGGAAGCCAG--CCAGGAGAGCAAGTTGTGTGGTCTATCCGTGAGGAC 2247
QY 2275 CRAAGGTGACCACTGGAAGGAGGCGCTGTCTGCTCCACAAGTCTCTGAACCTTTATCAG 2334
Db 2248 CAGGGCGGCGAGTGAAGACACGGCGGGATCATCTCCCGAGCTACGACATGGAGTACCAG 2307
QY 2335 GTGATTTTCGAGGGCGAAATCGGAAAAGGAAACCTTGGTGGATTTCTGTGGATGACATT 2394
Db 2308 ATTGTGTTCCGAGGAGTATAGGAAAGGACGTTCCGGAGAGATTGCCATTGATGACATT 2367
QY 2395 AGTATTAAATACACATTTTCAACAAGAAGATTGTGCAAAACCCAGCAGACCTGGATATAAAG 2454
Db 2368 CGGATAAGCACTGATGTCCCACTGGAGAACTGCATGGAACCC-----ATCTCG 2415
QY 2455 AACCCAGAATTAATTTGATGAACAGGAGGACGCCAGGATACGAAGTGAAGAGAA 2514
Db 2416 GCTTTTGCAGATGAATACAGAGTGGACTGGAGCAATCTTCTTCTCAACCTCAGGGTCT 2475
QY 2515 GGTGACAAGAATCTCCAGGAGCCAGGCAATGTGTTGAAGACCTTAGAACCCATCCCTC 2574
Db 2476 GGGGCC---CCTCGACCCGACAAAGAAAGAGTGCTGTACACCCCTGGATCCCATCCTC 2532
QY 2575 ATACCATATGACCATGAGCGCCCTGGGGTCTCTCTCTGGGGCTGTCTGT---GGGGTC 2631
Db 2533 ATACCATATGACCATGAGCTACCTGAGGCGTCTCTCTGGGGCCACCTGTGCGAGCCCTC 2592
QY 2632 GTGCTTACTGTCCCTGTTGGCATAATGGATGTCAGAAAGAACTTGTCTGCCCTGGAG 2691
Db 2593 CTGCTTACTGACCTGTTCTCTACTCGGCGCTGAGCTCCCGAAGCTGCACACACTGGAG 2652
QY 2692 AACTATACTTTGAACCTGTGGATGGTGTGAAGTTGAAAAAGACAACTGAATACACAG 2751
Db 2653 AACTACAACTTCGAGCTCTACGATGCGCTAA---GCACAAGTCAAGATGAACACCACAA 2709
QY 2752 AGTACTTATTCGGAGCATGA 2772
Db 2710 AAGTGCTGCTCCGAGGCATGA 2730

RESULT 9

US-08-936-135-13

; Sequence 13, Application US/08936135
; Patent No. 6054293

GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,135
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC97-288-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4769 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-936-135-13

Query Match 20.4%; Score 566.2; DB 3; Length 4769;

Best Local Similarity 54.1%; Pred. No. 4.6e-160;

Matches 1476; Conservative 0; Mismatches 1133; Indels 120; Gaps 11;

QY 107 CCGGGTACCTTACATCTCCTGGTTATCCTCATTCCTTATCACCCCAAGTGAATAATGGGAAT 166
Db 676 CTGGCTATACATCTCCCGAGGCTACCCCGAGACTATCCCTCCACGACACTGTGAGT 735
QY 167 GGTGATTCAGGCTCGGACCCATACAGAGAATTATGATCAACTTCAACCCCTCACTTCG 226
Db 736 GGATTCTACGCCCCGAAACCAACCAAGAGATTGTCTCAACTTCAACCCCTCACTTTG 795
QY 227 ATTTGAGGACAGAGACTGCAAGTATGACTACGTGGAAGTCTTCGATGGAGAAATGAAA 286
Db 796 AATTCAGAAACACGACTGCAAGTATGACTTCAATGAGATTCGGGATGGGACAGTGA 855
QY 287 ATGGACATTTTGGGAAAGTTCTGTGAAAGATAGCCCCCTCTCTCTGTGTCTTCTCAG 346
Db 856 CAGTCACCTCCTGGGAGCACTGTGGGAACATGCCCGCCACCATCATCTCCTCAG 915
QY 347 GGCATTTCTTTTATCAAAATTTGCTCTGACTACGAAACACATGTTGCGAGGATTTTCCA 406
Db 916 GTCCGTGTATACATCAAGTTTCACCTCAGACTACGCCCGGAGGGGGAGGTTTCTCTC 975
QY 407 TAGCTTATGAAATTTCAAGAGAGGTCCTGAA--TGTTCCCAAGAACTACACAAACACCTA 463
Db 976 TAGCTATGAGATCTTCAAAACAGGCTCTGAAGATTGTTCCAAAGAACTTTACAAAGCCCA 1035
QY 464 GTGGAGTATAAGTCCCCCGGATTCCTTGAAATAATCCCAACAGCCTTGAATGACATT 523
Db 1036 ATGGGACCATTTGAATCTCCAGGGTTTCCAGAGAAGTATCCACACAATCTGGAGTACCT 1095
QY 524 ATATTGCTTTGGCCAAAGATGTCAGAGATTATTCCTGGAATTTGAAGCTTTGACCTGG 583
Db 1096 TCACCATCTCGCCAAACCCAGGATGGAGTATCATCTACAGTTCCTTGACCTTTGACCTGG 1155
QY 584 AGCTGACTCAAAATCCTCCAGGGGGATGTCTGTCGCTACGACCGCTAGAAATCTGGG 643
Db 1156 AGCATGACCTCTTACAAGTGGGGGAAGGAGACTGTAATATGACTGGCTGGACATCTGGG 1215
QY 644 ATGGATTCCCTGATTTGGCCCTCACATTGGGGGTTTACTGTGGACAGAAACACAGGTC 703
Db 1216 ATGGCATCCACATGTTGGACCTCTGATTGGCAAGTACTGTGGGACGAAACACCTCCA 1275
QY 704 GAATCCGATCTCATCGGGCATTTCTCCATGGTTTTTACACCGACAGCGGATAGCAA 763
Db 1276 AACTCGCTCGTCCACGCGGATCCTCTCTCTGACCTTTTACACGAGCATGGCAGTGCCCA 1335
QY 764 AAGAAGTTTCTCAGCAAACTACAGTGTCTGCGAGGACAGTGTCTCAGAAAGATTTCAAAT 823
Db 1336 AGATGGCTTCTCCGCAACGTTTACTATTGATCCAGGAGGCCACCTGAGAAATTTTCAGT 1395

Db	2450	AAACTGCAGCTTTGAGGATGACAAAGATTGCAAACTCCTTCAGGATTCAAC-----	2501
Qy	1943	TTAACTGTGAATTTGGCTGGGGCTCTCAACAAGACCTCTGCACCTGGGAACATGACAATC	2002
Db	2502	-----TGCAACTTTGATTTCCGAGAGACCTGTGGTTGGGTGACGACCATG	2550
Qy	2003	ACGTGACAGCTCAAGTGGAGTGTGTACACAGCAAGACGGGACCCATTTCAGGATCACACAG	2062
Db	2551	CCAAAGTGGCTCCGGAGCAGGTGGATCAGCAGCGGCTTAACCCCAANTGACAAACATTTCCAC	2610
Qy	2063	GAGATGSCAACTTCATCTATTCCCAAGCTGACGAAAAATCAGAAGGGCAAGTGCTCGCC	2122
Db	2611	ATGACAAAGAACTTCTTGAACCTGCAGAGTGTATGGCCGACGAGAGGGCCAGTACGGCGGC	2670
Qy	2123	TGGTGAAGCCTGTGGTTTATTCCCAGAACTCTGCCCACTGCATGACCTTTCTGGTATCAC	2182
Db	2671	TCATCAGCCCCACGGTGCACCTGCCCCGAAGCCCTGTGTGCATGGAGTTCGAGTACCAAG	2730
Qy	2183	TGCTGGTCCCAGCTCGGCACACTCAGGTCACAGGTCGCTACCAAGAGCCACAGGAGT	2242
Db	2731	CCATGGGGCCGACAGGGGTGGCACTGCAGGT-----GGTTCGGAGGCCACCCAGGA	2782
Qy	2243	ACGATCAGCTGTGTGGATGGCCATTGGACACCAAGGTGACCACTGGAAGGAAGGGCGTG	2302
Db	2783	ANGCA-AACTCCTTTGGGTATCCGTGAGGACCGAGCCAGGGCAGCGTGGGAAGCAGGGCGCA	2841
Qy	2303	TCATTGTCACAAAGTCTCTGAACTTTATFCAGGTGATTTTCGAGGGCGAAATCGGAAAG	2362
Db	2842	TTATCCTGCCAGCTATGACATGGAGTATFCAGATCGTGTTCGAGGGAGTGTAGGGAAGG	2901
Qy	2363	GAACCTTGTGTGGATTTGCTGTGGATGACATAGTATTAAATACCAACATTTTCACAAGAG	2422
Db	2902	GACATCGGGAGAGATTTCGGGGATGACATTCGGATAGCACTGATGCCCATCTGGAGA	2961
Qy	2423	ATTGTGCAAAACACGACAGACCTGGATATAAAGAACCCAGAAATTTAAATTTGATGAAACAG	2482
Db	2962	ACTGCATGGAACCCATATCAGCTTTTTCAGTGGACATCCAGAAACCCATGGGGAGAGG	3021
Qy	2483	GGAGCAGCCAGATACGAAGGTGAAGGAGAGAGGTGACAAAGACATCTCCA-----	2533
Db	3022	GCTATGAAGATGAGATTGATGTAATATGAAGGAGATTGGAGCAACTCTCTTCTCCTCTA	3081
Qy	2534	-----GGAAGCCAGGCAATGTGTTGAAGACCTTAGAAC	2566
Db	3082	CCTCAGGGGTGTGACCCCTCATCTGGCAAAAGAGAGCTGGCTGTACACCTAGATC	3141
Qy	2567	CCATCTCATCACCATATAGCCATAGAGCGCCCTGGGGTCTCTCTGGGGGCTGTCTGTG	2626
Db	3142	CCATTCTGATCACCATCATCGCCATGAGCTCGCTGGGGTCTCTGTCTGGGGCCACCTGTG	3201
Qy	2627	GGG---TCGTGCTGTACTGTGCCTGTGGCATTAATGGGATGTCAGAAAGAACTTGTCTG	2683
Db	3202	CGGGCTCTCTCTTACTGCACCTGCTCTCTATTCGGGTCTGAGTTCGAGGAGCTGCACCA	3261
Qy	2684	CCCTGGAGAACTATACTTTGAACCTTGTGGATGGTGTGAAGTTGAAAAAGACAACTGA	2743
Db	3262	CAGTGGAGAACTACAACTTTGACCTACGATGGCCCTCA---GCACAAGTCAAGATCA	3318
Qy	2744	ATACACAGAGTACTTATTCGGAGGCATGA	2772
Db	3319	ATCATCAGAAAGTGTCTGCGGAGGATGA	3347

```

RESULT 10
US-09-116-473-1
: Sequence 1, Application US/09116473
: Patent No. 6428965
: GENERAL INFORMATION:
: APPLICANT: Kolodkin, Alex
: APPLICANT: Ginty, David
: TITLE OF INVENTION: SEMAPHORIN RECEPTOR
: NUMBER OF SEQUENCES: 8

```


QY	1675	ACGCGATTTCATCAGGATCTTACCCCGAGAGAGCCACTCATGCGGAGCTTGGGCTCAGAAATG	1734
Db	2013	GCACAGTACGTGCGGGTATACCCAGAGAGGTGGTCACCAGCGGGCATCGGGATGAGCGTG	2072
QY	1735	GAGTGTCTGGGCTGTGAACTGGAAGCCCCCTACAGCTGTGACCGACCCACTCCCAACGGGAAC	1794
Db	2073	GAGGTCTCTGGGCTGTGACTGTGGACAGACTCAAAGCCCCACAGTGGAGACG	2128
QY	1795	TTGGTGGATGAATGTGATGACGACGAGGCCAACTGCCACAGTGGACAGAGTGATGACTTC	1854
Db	2129	CACCGTAAGAGTGAAGAGACCAACCACCACATATCCCATGGATGAGGATGCCACGGAGTG	2188
QY	1855	CAGCTCACAGGTGGCACCACTGTCTGSCCAGACAAAAGCCCCCGGTATACAGACGACCC	1914
Db	2189	TGGGGAAAA--CTGCGAGCTTTGAGGATGACAAAGATTTGCAACTTCTCTCAGGATTCAAC	2246
QY	1915	ATACAATCAGAGTTTCCACACATATGGTTTAACTGTGAATTTGGCTGGGGCTCTCAACAG	1974
Db	2247	TGCAACTTGTATTTCTTGAGNA-----	2269
QY	1975	ACCTTCTGCCACTGGGAACATGACAATACGCTGCAGCTCAAGTGGAGTGTGTTGACACGC	2034
Db	2270	--GACCTGTGTTGGATGTACGACCGTGCCAAGTGGCTGCAGAGTACCTGGATCAGCAGT	2327
QY	2035	AAGACGGGACCCATTTCAGGATCACAGGAGATGCAACTTCATCTATTCCCAAGCTGAC	2094
Db	2328	GCCAAACCTTAACGACAGAACGTTTCCAGATGACAAAGAACTTCTTCAAACATACAGAGCGAC	2387
QY	2095	GAAATCAGAAAGGCAAGTGGCTGCCTGGTGGAGCCCTCTGGTTTATTTCACAGAACTCT	2154
Db	2388	GGCGGAGGAGGGCCAGTTTGGCGGCTCATCAGCCCAACAGTGCACCTGCCCGCAAGC	2447
QY	2155	GCCCACTGCATGACCTTCTGGTATCACATGTCTGGGTCCCACGCTCGGCACACTCAGGGTC	2214
Db	2448	CCTGTGTGATGTAGTTTCCAATACCAAGCCATGGCGGCCACAGGGTGGCACTGCAGGT-	2506
QY	2215	AAACTGGCTTACCAGAAGCCAGAGGAGTACGATCAGCTGGTCTGGATGGCCATTGGACAC	2274
Db	2507	-----GGTTCGGAAGCCAGACAGGAAGC--AAACTTCTTTGGGTCTATCCCGGAGGAC	2558
QY	2275	CAAGGTGACCACCTGGAAGGAAGGGGTGCTTGTCTCCACAAAGTCTCTGAAACTTTATCAG	2334
Db	2559	CAGGCGAGGAGTGAAGCATGGACGCATTATCTTGCCTCCGCTATGACATGGAGTATCAG	2618
QY	2335	GTGATTTTCGAGGGCGAAATTCGAAAGAAGAAACCTTTGTGGGATGTCTGGATGACATT	2394
Db	2619	ATCGTATTCGAGGGAGTGTATCGGAAAGGGCGATCCGAGAGATTTCCATCGACGACATT	2678
QY	2395	AGTATTAATAAACACATTTTCAAGNAGATTCTGCAAAACCCAGACACCTGGATAAAAG	2454
Db	2679	CGGATAGCACCCGATGTCCCATCTGGAGAACTGCATGGAACCCCATCTCGGCTTTTGCAGTG	2738
QY	2455	AACCCAGAAATTAATAATTGATGAAACAGGGAGCGCCAGGATACGAAGGTGAAGAGAA	2514
Db	2739	GACATCCAGAAATCCATGGGGAGAGSGCTATGAAGATGAGATTGATGACTATGAA	2798
QY	2515	GGTGACNAGAACATCTCCA-----	2541
Db	2799	GGAGATTGGAAACAACCTTCTTCTCTACCTCAGGGGTGTAGTGCCTCATCTGGCAAGAA	2858
QY	2542	GGCAATGTGTTGAAGACCTTAGAACCCATCTCATCACCATCATTAGCCATGAGCGCCCTG	2601
Db	2859	AAGAGTGGCTGTACACACTGACCCCATCTTGATACCCATCATTTGCCATGAGCTCGCTG	2918
QY	2602	GGGGTCTCTCTGGGGCTCTGTGGGG---TCGTGCTGTACTGTGCTCTTGGCATTAAT	2658
Db	2919	GGTGTCTGTCTGGGGCCACCTGTGGGGCCCTCTCTCTACTGCACCTGCTCTACTCT	2978
QY	2659	GGGATGTCAGAAAGAACTTGTCTGCCCTGGAGAACTATAACTTTGAACTTGTGGATGGT	2718
Db	2979	GGCCTGAGTTTCGAGAGCTGCACCACTACCTGGAGAACTACAACCTTTGAGCTCTACGACGGC	3038

```

QY 2719 GTGAGTTGAAAAAAGACAAATGTAATACACAGAGTACTTATTCGAGGCGCATGA 2772
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3039 CTCAA---GCACAAGTCAAGATCAATCACCAGAAAGTGCTGCTCGAGGCGCATGA 3089
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-08-936-135-19
; Sequence 19, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936.135

```

Query Match	20.3%	Score 561.8;	DB 3;	Length 2781;
Best Local Similarity	54.0%	Pred. No. 7e-159;		
Matches 1490;	Conservative	0;	Mismatches 1147;	Indels 120;
Gaps				

Qy	79	TGTGGCGATACTATAAAATTGAAAGCCCGGTACTTTACATCTCTCTGGTTATCTCAT	138
Db	82	TGCGGGGTGTTTGAATTCCTCAAAAGATGCTGGGTATATCACTCTCCGGTTACCCCCAG	141
Qy	139	TCATTATCACCAAGTCGAAAAATGCGAATGCTGATTTCAGGCTCCGGACCCATACCCAGA	198
Db	142	GACTACCCCTCCACACAGAACTGGAGTGGATTGTTCGCCCCCGAACCCACCCAGAAG	201
Qy	199	ATTATGATCAACTTCAACCCCTCACTTCGATTTCGGAGCAGAGACTGCAAGTATGACTAC	258
Db	202	ATTGTCCTCAACTTCAACCCCTCACTTTGAAATCGAGAAGCACCACTGCAAGTATGACTTT	261
Qy	259	GTGGAAGCTTCGATCGAGAAAATGAAATGGACATTTTAGGGGAAAGTTCTGTGGAAG	318
Db	262	ATCGAGATTCGGATGGGACAGTGAATCCGACAGACTCCTGGGCAACACTGTGGGAAC	321
Qy	319	ATAGCCCTCCTCCTGTGTGCTTCAGGGCCATTTCTTTTATCAAAATTTGTCCTGTGAC	378
Db	322	ATGCCCCGCCACCATCATCTCTCTCGGGTCCATGCTCTACATCAAGTTCCACCTCCGAC	381
Qy	379	TACGAAACACATGGTCAGGATTTTCCATACGTTATGAAATTTTCAAGAGAGGTCCTCGAA	438

Db 382 TAGCCCGGCGAGGGGCGAGCTTCTCTCTGCGCTACGAGATCTTCAAGACAGGCTCTGAA 441
QY 439 ---TGTTCOCAGACTACACACACCTAGTGGAGTGATAAAGTCCCGGATTCCTCGAA 495
Db 442 GATTGCTCAAAAACCTTCAAGCCCAACAGGGACCATCGAATCTCTGGGTTTCTGAG 501
QY 496 AATATATCCAAACGCTTGAATGCATTAATTTGCTTGGCGCAAGATGTCAGAGATT 555
Db 502 AAGTATCCACACAACTTGGACTGCACTTACCATCTGCGCCAAACCCAAAGATGGAGATC 561
QY 556 ATCTTGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAACTCTCCAGGGGGAATGTC 615
Db 562 ATCTCGAGTTCCTGATCTTTGACCTGGAGCATGACCCCTTGCAGTGGGAGGGGAC 621
QY 616 TGTGCTAGACCGGCTAGAAATCTGGGATGGATTCCTCGATGTTGGCCCTCACATTTGGG 675
Db 622 TGCAAGTAGGATTTGGTGGACATCTGGATGGCAATCCACATGTTGGCCCTGATTTGGC 681
QY 676 CGTTACTTGGACAGAAACACACAGGTGGAATCCGATCCTCATCGGGCAATTTCTCCATG 735
Db 682 AAGTACTGTGGGACCAAAACACCTCTGAATTCGTTTCATCGACGGGATCTCTCCCTG 741
QY 736 GTTTTTTACACGACAGCGGATAGCAAAAGAGGTTTCTCAGCAAACTACAGTGTCTTG 795
Db 742 ACCTTTACACGACATGGCGTGGCCAAAGGATGGCTTCTCGCGCTTACTACCTGGTC 801
QY 796 CAGAGCAGTGTCTCAGAAGATTTCAATGTATGGAAGCTCTGGGCATGGAATCAGAGAA 855
Db 802 CACCAAGAGCCACTAGAGAATTTTCAAGTCAATGTTCTCTGGCATGGAGTCTGSCCG 861
QY 856 ATTCATCTGACAGATCACAGTTCTT---CCAGTATAGCAACAACTGGTCTGCAGAG 912
Db 862 ATTGCTAATGAACAGATAGTGCCTCATCTACTCTGATGGGAGTGGACCCCTCAA 921
QY 913 CGTCTCCGCTGAATACCTGAGATGGTGGACTCCGGAGAGGATTCCTACCCAGAG 972
Db 922 CAAAGCGGCTCATGGTGATGCAATGGCTGGACCCCAACTTGGATTTCAACAAGGAG 981
QY 973 TGGATACAGTAGACTTGGGCTTCTGCGCTTTGTGTCACGGCTGTGCGGACACAGGCGGC 1032
Db 982 TATCTCCAGTGAOCTGCGCTTTTAAACATGCTCAAGCCCATCGCAACACAGGAGCG 1041
QY 1033 ATTTCAAAAGAACCAAGAGAAATATATGCAAGCTTCAAGATCGACGTTAGCTCC 1092
Db 1042 ATTTCCAGGGAACACAGAAATGSGCTACTACGTCAAACTCTACAAGCTGGAAGTCAGCACT 1101
QY 1093 AACGGGAAGACTTGGATACCATATAAAGAGGAACAAACCTGTTCTCTTTCAGGGAAC 1152
Db 1102 AATGGAGAGCTGGATGGTGATCCGGCATGGCAAAACCAACAGGTATTTCAAGCCAAC 1161
QY 1153 ACCAACCCACAGATGTTGTTGAGTATTTCCCAAAACCACTGATAACTCGATTGTC 1212
Db 1162 AAGGATGCAACTGAGTGGTCTTGAACAAGCTCCACGCTCCACTGTCGACAGGTTGTT 1221
QY 1213 CGAATCAAGCTGCAACTTGGGAACTGGCAATCATATGAGATTTGAATATACGTTGTC 1272
Db 1222 AGAATCCGCGCTCAGACCTGGCACTCAGGTATCGCCTCCGCTGGAGCTCTTGGGCTGC 1281
QY 1273 AAGATAACAGATATCTTGTCTGGAATGTTGGGTATGGTGTCTGGACTTATTTCTGAC 1332
Db 1282 CGGGTCAAGATGCTCCCTGCTCAACATGCTGGGATGCTCTCAGGCGCTCATTCGACAG 1341
QY 1333 TCCAGATCACATCATCCAAACCAAGAGACAGAAATGATGCTGAAAACATCCGCGCTG 1392
Db 1342 TCCAGATCTCGGCTCTTCCACCCAGGAATACCTTGGAGCCCAAGTCCAGCCCGCTG 1401
QY 1393 GTAACAGCTCGTCTGGCTGGGACATTCOACCCGACCTCATTTCTTACATCAATGAGTGG 1452
Db 1402 GTCAGAGCGCGCTGGGCTGCTTCCCTCGAATCCCTCAGGCCCAAGCCCGGTGAGGAGTGG 1461
QY 1453 CTCCAATAGACTGGGAGAGAGATCGTGGGGGATCATCATTCAGGGTGGGAG 1512
Db 1462 CTTCAAGGTAGATCTGGGAACACCCCAAGACAGTGAAGGTTGTATCATCCAGGAGCCCGC 1521

QY 1513 CACCCGAGAG-----ACAAGTGTTCATGAGGAAGTTCAAGATCGG 1554
Db 1522 GGAGGAGACAGTATCACTGCTGTGGAAGCCAGAGCATTTGTGCGCAAGTTCAAGTCTCC 1581
QY 1555 TACAGCAACAAACGGCTCGGACTGGAAGATGATCATCGATGACAGCAACCAAGGCGAAG 1614
Db 1582 TACAGCTTAACGGCAAGACTGGGAATACATTCAGGACCCCAAGGACCCAGCAAGCAAG 1641
QY 1615 TCTTTTGGGGCAACAACTATGATACACCTGAGCTGCGGACTTTTCCAGCTCTCTCC 1674
Db 1642 CTGTTCAAGGAAACATGCACTATGACACCCCTGACATCCGAAGGTTTGACCCCATCCG 1701
QY 1675 ACBCGATTCATCAGGATCTACCCGAGAGAGCCACTCATGCGGACTTGGGCTCAGAAATG 1734
Db 1702 GCACAGTATGTGCGGGTATATCCGGAGAGGTGTGTCGCCGCGGGGATTTGGATGCGGCTG 1761
QY 1735 GAGCTCTCGGCTGCAAGTGGAAAGCCCTACAGCTGGAGCCGACACTCTCCAAACGCGAAG 1794
Db 1762 GAGGTCTCGGCTGAGCTGGAGACTCCAAAGCCCAAGCGTAAACACG-----TGGGACC 1817
QY 1795 TTTGGTGAATGATGATGACGACAGGCAACTGCCACAGTGGAAACAGTGTGATGACTTC 1854
Db 1818 CACTGTGAAGGGAAGAGACAAACACCCCTACCCCAAGAGAGAGAGGCA-----1870
QY 1855 CAGCTCACAGGTGGCACCACTGTGCTGGCCACAGAAAAGCCCAAGGCTCATAGACAGCACC 1914
Db 1871 CAGAGTGTGGGAGAACTGCAGCTTTGAGGATGACAAAG-----ATTTGCACTCC 1921
QY 1915 ATCAATCAGAGTTTCCAAACATATGTTTTAACTGTGAATTTGGCTGGGCTCTCACAAG 1974
Db 1922 CTTGCGGATTTCAATTGCA-----CTTCGATTTCTTCGAG 1956
QY 1975 ACCTTCTGCACTGGGAACATGACATCACGTGCAGCTCAAGTGGAGTGTGTGACACG 2034
Db 1957 GAGCCTGTGTTGGATGTATGACCATGCCAAGTGGCTCCGACCACTTGGGCCAGCAGC 2016
QY 2035 AAGACGGGACCCATTCAAGATCACACAGAGATGGCAACTTCATCTATTCCCAAGGTGAC 2094
Db 2017 TCCAGCCCAACAGCAGCGGAGCTTCCAGATGACAGGAATTTCTTGGGCTGCAGAGTGAC 2076
QY 2095 GAAATCAGAAGGGCAAGTGGCTGCCCTGGTGGAGCCCTGTGGTTTATTTCCCAAGACTCT 2154
Db 2077 AGCCAGAGAGAGGGCGAGTATCCCGGCTCATCAGCCCTCTCCACCTGCCCCGAAGC 2136
QY 2155 GCCCACTGATGACCTTCTGGTATCATATGCTGGGTCCACGTGCGGACACTCAGGGTCT 2214
Db 2137 CCGGTGTGATGGAGTTCCAGTACAGGCCACGGGCGGCGGGGTGCGGCTGCAGGT- 2195
QY 2215 AAATGCGCTACCAAGGCCAGAGAGTACGATCAGCTGGTCTGGATGGCCATTGGACAC 2274
Db 2196 -----GTTGCGGGAAGCCAGCC-AGGAGAGCAAGTTCTGTGGGTCTATCCGTGAGGAC 2247
QY 2275 CAAGTGACCACTGGAGGAAGGGGTGTCTTGTCTCCCAAGTCTCTGAACTTTATCATG 2334
Db 2248 CAGGGCGGAGTGGAAACAGCAGCGGCGATCATCTGCCAGCTACGACATGGAGTACCA 2307
QY 2335 GTGATTTTCAGGGGGAATCGAAAAAGAAACCTTGTGGGATGCTCTGTGATGACATT 2394
Db 2308 ATTGTTTCAGGGGAGTGATAGGGAAGGACGTTCCGGAGAGATTGCCATTGATGACATT 2367
QY 2395 AGTATTAATAACCACTTTCAACAAGAAGATTGTGAAAACACAGCAGACCTTGATAAAAAG 2454
Db 2368 CGGATAAGCACTGATGTCCCACTGGGAGACTGCATGGAACTCTCGGCTTTTTCAGTG 2427
QY 2455 AACCCAGAAATTAATTTGATGAAACAGGGAGCAGCCAGGATACAGGTGAAGAGAA 2514
Db 2428 GACATCCCAAGAAATACATGAGAGAAAGGATATGAAGATGAAATGATGATGAATACGAG 2487
QY 2515 GGTGACAGAAACATCTCCAGGAGGAGCCAGCAATGTCT-----2551
Db 2488 GTGGAGTGAGCAATTTCTTCTTCTGCAACCTCAGGGGTCTGGCGCCCTCTCGACCGACAAA 2547

Qy	1361	ACAGAACTGGATGCCGTGAAACATCCGGCTGGTAAACAGTCGTCTGGCTGGGCATCTC	1420
Db	1936	AGTACCTCTGGAGCCCGAGCTGCTGCCCGCTGGTTAGTAGCGGCTCTGGCTGGTTCTC	1995
Qy	1421	CACCCGACCTCATCTCATCAATGAGTGCCTCCAATAGACTCGGGGAGGAGAGA	1480
Db	1996	GGAACTCAAGCCCGACCCAGGTGAAGATGCTTCAGTGACTCGGGACACCCAAGA	2055
Qy	1481	TCGTAGGGGCATCATCTACAGGTGGGAACACCGGAGA	1522
Db	2056	CAGTGAAGGGGTCACTATCCAGGAGCCGAGGAGGAGACAGCATCACTCCCGTGAAG	2115
Qy	1523	ACAAGGTGTCATGAGGAAGTTCAGATCGGTCAGACAACAACGGCTCGACTGGAAGA	1582
Db	2116	CCAGGGGTTTGTACGCAAGTTCAAAAGTCTCTACAGCTAAATGGCAAGACTGGGAAT	2175
Qy	1583	TGATCATGGATGACAGCAAAACGCAAGGGCAAGTCTTTGAGGGCAACAACAATCATGATA	1642
Db	2176	ATATCCAGGACCCAGGACTACGACAGCAAAAGCTGTTGAAGGNAACATGCATATGACA	2235
Qy	1643	CACCTGAGCTGCGGACTTTTCCAGCTCTCTCCACCGATTCATCAGGATCTACCCGAGA	1702
Db	2236	CCCTGACATCCGAAGTTTCATCTGTTCCAGCGCAGTATGTGCGGCTGTACCCAGAGA	2295
Qy	1703	GAGCCATCATGCGGACTGGGGCTCAGAAATGGAGCTGCTGGGCTGTGAAGTGAACCC	1762
Db	2296	GGTGGTCGCCAGCAGGATCGGGGATGAGCTGGAGGTGCTGGGCTGTACTGGACAGACT	2355
Qy	1763	CTACAGCTGGACCGACCACTCCCAACGGGAACCTGGTGGATGAATGTGATGACGACCCAGG	1822
Db	2356	CAAGCCACAGTGGAGAGCGCT---GGAGCCACCGTGAAGAGTGAAGAGACTACCA	2411
Qy	1823	CCAACTGCCAGTGGAAAGGTGATGACTTCCAGCTCACAGTGGCACCACTGTGCTGG	1882
Db	2412	CCATATCCCATGGATGAGATGCCA-----CCGAGTGTGGGA	2449
Qy	1883	CCACAGAAAGCCACGGTCATAGACAGCACCATACATCAGATTTCCAACTATGGTT	1942
Db	2450	AACTGCAGCTTTGGAGATGACAAGATTGCAACTCTCTTCAGGATTCAC	2501
Qy	1943	TTAACTGTAATTTGGCTGGGCTCTCAACAGACCTTCTGCCACTGGGAACATGACAATC	2002
Db	2502	-----TGCAACTTTGATTTCCGGAAGACCTGTGGTGGTGTACGACCATG	2550
Qy	2003	ACGTGCAGCTCAAGTGGAGTGTGTACACAGACAGCGGACCATTCAGGATCACACAG	2062
Db	2551	CCAACTGGCTCCGGAGACGCTGGATCAGCAGCGCTTAACCCCAATGACAGAACTTCCAG	2610
Qy	2063	GAGATGCAACTTCATCTATCCCAAGCTGAGGAAATCAAGAGGCAAACTGGCTGCC	2122
Db	2611	ATGACAAGAACTTCTTGAACCTGCAGAGTGAATGGCCGACGAGAGGGCCAGTACGGGGCC	2670
Qy	2123	TGCTGAGCCCTGTGGTTATTCCCAAGACTCTGCCCACTGCATGACCTTTGTTATCACA	2182
Db	2671	TCATCAGCCACCGGTGCACCTGCCCGAAGCCCTGTGTGATGGAGTTCAGTACCAAG	2730
Qy	2183	TGCTGGTCCACGTGCGGACACTCAGGGTCAAACTCGGCTACGAAAGCCAGAGGAGT	2242
Db	2731	CCATGGCGGGCACGGGTGGCACTCGAGT-----GGTTCCGGAAGCCAGCCAGGA	2782
Qy	2243	ACGATCAGCTGGTCTGGATGGCCATTGGACACCAAGGTGACCTGGAAGGAGGGCGTG	2302
Db	2783	AAGC-AACTTCTTTGGGTATCCGTGAGGACCGAGGCGAGGTGGAGACCGGGCGCA	2841
Qy	2303	TCTTGCTCCCAAGTCTCTGAACTTTATCAGGTGATTTTCGAGGGGGAATCGGAAAG	2362
Db	2842	TTATCTGCCAGCTATGACATGGAGTATCAGATCGTGTTCGAGGAGTGTATAGGGAAG	2901
Qy	2363	GAACCTTGGTGGGATGCTGTGGATGACATAGTATTATAACCACTTTACAGAAG	2422
Db	2902	GAGCTCGGGAGAGATTTCCGGCATGACATTCGGATGAAGCACTGATGTCTCCCATGAGA	2961

QY	2423	ATTGTGCAAAACC----	AGCAGACCTCGGATAAAAGAACCCAGAAATTTAAATTTGATGAAA	2479
DB	2962	ACTGCATGGAACCCATATCAT	CAGCTTTTTCGAGGTGAGGATTTTAAAGTGGACATCCAGAAA	3021
QY	2480	CAGGGAGCACGCCAGGATACGA	AGGTGAAGGAGAAAGGTGACAAGAA-----	2525
DB	3022	CCATGGGGAGAGGGCTATGA	AGATGAGATTGATGATGAATATGAGAGGAGATTGGAGCA	3081
QY	2526	-----	-----CATCTCCAGGAAGCCAGGCAATGTGT	2551
DB	3082	ACTCTTCTCTCTACTCAT	AGGGGCTGGTGACCCCTCATCTGGCAAAAGAAAGAGCTGGC	3141
QY	2552	TGAAGACCTTTAGAACCCATC	CATCACCATCATAGCCATGAGCGCCCTGGGGGTCTCTCC	2611
DB	3142	TGTACACCTTAGATCCCATCT	GATCACCATCATCGCCATGAGCTGCGTGGGGGTCTCTGC	3201
QY	2612	TGGGGGCTGTCTGTGGG---	TCGTCTCTACTGTGCTGTGGCATAATGGGATGTGAG	2668
DB	3202	TGGGGGCCACTGTGGGGCT	CTCTCTTACTTGACCTGCTCTTATTCGGGTCTGAGTT	3261
QY	2669	AAAGAAACTTGTCTGCCCT	GGAGAACTAATCTTGAACCTTGTGGATGGTGTGAAGTTGA	2728
DB	3262	CGAGGAGCTCACCACTGG	AGAACTCAACTTTGAGCTCTACGATGGCCTCAA---GC	3318
QY	2729	AAAAGACAAACTGAATAC	CACAGAGTACTTATTCGGAGGCATGA	2772
DB	3319	ACAAGGTCAAAGATCAAT	CATCAGAAGTGTGCTCGGAGGCATGA	3362

RESULT 13
 US-08-936-135-21
 ; Sequence 21, Application US/08936135
 ; Patent No. 6054293
 ; GENERAL INFORMATION:
 ; APPLICANT: Tessier-Lavigne, Marc
 ; APPLICANT: He, Zhigang
 ; APPLICANT: Chen, Hang
 ; TITLE OF INVENTION: Semaphorin Receptors
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 75 DENISE DRIVE
 ; CITY: HILLSBOROUGH
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94010
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/936,135
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OSMAN, RICHARD A
 ; REGISTRATION NUMBER: 36,627
 ; REFERENCE/DOCKET NUMBER: UC97-288-2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 343-4341
 ; TELEFAX: (650) 343-4342
 ; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4765 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 US-08-936-135-21

Query Match 19.4%; Score 537.8; DB 3; Length 4765;
Best Local Similarity 58.9%; Pred. No. 1.7e-151;

Matches	992;	Conservative	0;	Mismatches	667;	Indels	24;	Gaps	3;
---------	------	--------------	----	------------	------	--------	-----	------	----

QY	107	CCGGTAGCTTACATCTCTCGTGTATCTCCATCTTATCACCCCAAGTAAAAATGCGAAT	166
DB	676	CTGGCTACATCACTTCCCCAGGCTACCCCCAGGACTATCCCTCCACAGAACTGTGAGT	735
QY	167	GGCTGATTGAGGTCGCGGACCCATACCAGAGAAATTTATGATCAACTTCAACCTCACACTTCG	226
DB	736	GGATTGTCTACGCCCCCGAAACCCACCAAGAGATTGTTCTCACTTCAACCTCACACTTCG	795
QY	227	ATTTGGAGCAGACAGACTCGAAGTAGACTACGTGGGAAGTCTTCGATGGAGAAAATGAAA	286
DB	796	AAATCAGAAACACGACTCGAAGTAGACTCATTTGAGATTGGGATTCGGGATGGGACAGTAGT	855
QY	287	ATGGACATTTTAGGGGAAAGTTCTGTGGAAAGATAGCCCCCTCCCTGTTGTGTCTTCAG	346
DB	856	CAGCTGACCTCTGGGCAAGCACTGTGGGAACATTCGCCCCGCCACCAATCATCTCCCTCAG	915
QY	347	GGCCATTTCTTTTATCAAAATTTGTCTGTGACTAGCAAAACACATGGTGCAGGATTTTCCA	406
DB	916	GCTCCGTGTATACATCAAGTTTCACTCAGACTACGCCCGGACGGGCGAGTTTCTCTC	975
QY	407	TAGCTTATGAATTTTCAAGACAGGTCTCTGAA--TGTTCCAGAACTACACAAACACTA	463
DB	976	TAGCTTATGAGATCTTCAAACAGGCTCTGAAGATTGTTCCAAGAACTTTACAAGCCCA	1035
QY	464	GTGGAGTGATAAAGTCCCCCGGATTCCCTGAAAAATATCCCAACAGCCTTCAATGCACTT	523
DB	1036	ATGGGACCATTGAACTCCAGGGTTTCCAGAGAAGTATCCACAATCTGGACTGTACTCT	1095
QY	524	ATATTGTCTTTGGGCAAAAGATGTCCAGAGATTATCCTTGGAAATTTGAAGCTTTGACCTGG	583
DB	1096	TCACCATCTGGCCAAACCCAGGATGGAGTATCCTACAGTCTCTGACCTTTCGACCTGG	1155
QY	584	AGCCTGACTCAAAATCCTCCAGGGGGATGTTCTGTGCTACGACGGGTAGAAATCTGGG	643
DB	1156	AGCATGACCTCTACAAAGTGGGGGAAGGAGACTGTAAATATGACTTGGGTGACATCTGGG	1215
QY	644	ATGGATTCCCTGATGTGGCCCTCAATTTGGGCGTTACTGTGGACAGAAACACACAGTCT	703
DB	1216	ATGGCATTCACATGTTGGACCTCTGATTTGGCAAGTACTGTGGACGAAACACCCCTCCA	1275
QY	704	GAATCCGATCTCATGCGGCATCTCTCCATGGTTTTTACACGACGACGCGCATAGCAA	763
DB	1276	AACCTCGCTGCTCCACGGGGATCTCTCTCTGACCTTTCACACGGACATGSCAGTGGCCA	1335
QY	764	AACAAGTTTCTCAGCAAACTACAGTGTCTGACAGACAGTGTCTCAGAGAATTTCAAT	823
DB	1336	AGGATGGCTTCTCCGACGTTTACTATTGTATCCACGAGGACCCACCTTGAAATTTTCAGT	1395
QY	824	GTATGGAAGCTCTGGGCATGGAATCAGGAGAAATTCATCTGACACAGATCACAGCTTCTT	883
DB	1396	GCAATGCTCCTTTGGGAATGGAGTCTGCGCGGATTGCTAATGACAGATCAGTGCCTCT	1455
QY	884	---CCAGTATAGCACCAACTGGTCTGCAGAGCGCTCCCGCTTGAACCTACCCCTGAGAATG	940
DB	1456	CCACCTTCTCTGATGGAGGTGGACTCTCAACAGAGCGGCTCCATGGTGTGACAAATG	1515
QY	941	GGTGGACTCCCGAGAGGATTCCTACCGAGAGTGGATACAGGTAGACTTGGGCCCTTCGC	1000
DB	1516	GCTGGACACCCAAATTTGGATTCCAAAGAGGAGTATCTCCAGGTGGACCTGCGCTTCCTAA	1575
QY	1001	GCPTTCTCACGGCTGTGGGACACAGGGCGCATTTCAAAAGAAACCAAGAAAGAAATTT	1060
DB	1576	CCATGCTCACGCCATTGCAACACAGGAGGCCATTTCCAGGGAACCCAGAAAGGCTACT	1635
QY	1061	ATGTCGAAGACTTACAAGATTCGAGTTAGCTCCAACGGGGAAGACTGGATCACCATAAAG	1120
DB	1636	ACGTCAAAATCGTACAAGCTGGAAGTCAGCAAAATGGTGAAGATTGGATGGTCTACCGGC	1695
QY	1121	AAGGAACAAACCTGTTCTCTTTTCAGGGAACACCAACCCACAGATGTTGTGGTGTGCAG	1180
DB	1696	ATGGCAAAACACAAAGATATTCCAAAGCGAAACATGATGCCACCGAGTGTGTCTAAACA	1755

Qy	1181	TATTCCCAACACATGATAACTCGATTGTGTCGGAATCAAGCCTGCAACTTTGGAAACTG	124
Db	1756	AGTCCACATGCCACTGCTGACTCGGTTCATCAGGATCCGCCGACAGCTGGCATTTGG	1815
Qy	1241	GCATATCTATGAGATTTGAAGTATACGGTTGCAAGATAACAGATTATCTTTGCTCTGGAA	1300
Db	1816	GCATTGCCCTTCCCTCTGGAGCTCTTTGGCTGCGCGGTACACAGATGCACCTCTCCAACA	1875
Qy	1301	TGTTGGGTATGGTCTGCGACATATTCTTCGACTCCACATCACATCATCAACCAAGGAG	1360
Db	1876	TGCTGGGGATGCTCTCGGGCTCATTTGTGATACCCAGATCTCTGCCCTCCACCCGAG	1935
Qy	1361	ACAGAACTGGATGCCCTGAAAACATCCGCCCTGGTAACCAAGTCGTCTGGCTGGGCATTC	1420
Db	1936	AGTACCTCTGGAGCCCCACGTGCTGCCGCCCTGGTTAGTAGCCGCTCTGGCTGGTTTCCTC	1995
Qy	1421	CACCCGACCTCATTTCTTACATCAATGAGTGGCTCCAAATAGACTTGGGGGAGGAGAAG	1480
Db	1996	GGAACTCTAAGCCACAGCCAGGTGAAGATGGCTTCAGGTAGACCTGGGGACACCCAAGA	2055
Qy	1481	TCGTGAGGGGCATCATCTTACAGGTGGGAAGCAGCCAGAGAACAA-----	1526
Db	2056	CAGTGAAGGGGTTCATCTCCAGAGAGCCGAGGAGGACAGCATCACTGCCGTGGAAG	2115
Qy	1527	----GGTGTTCATGAGGAAGTTCAAGATCGGGTACAGCAACACGGCTCGGACTGGAAGA	1582
Db	2116	CCAGGGCGTTTGTACGCAAGTTCAAGTCTCTACAGCTTAATGGCAAGGACTGGGAAT	2175
Qy	1583	TGATCATGATGACAGCAACCGAAGCGAAGTCTTTTTCGAGGGCAACACAACTATGATA	1642
Db	2176	ATATCCAGGACCCAGGACTACGACAGCAAAAGCTGTTTGAAGGAACATGCATCATGACA	2235
Qy	1643	CACCTGAGCTGGGCACTTTTCCAGCTCTCTCCACCGCATTCATCAGAGATCTACCCCGAGA	1702
Db	2236	CCCTTGACATCCGAAGTTTCGATCTGTTCCAGCCGACGTATGTGCGGGTGTACCCACAGA	2295
Qy	1703	GAGCCACTCATGGCGACTTGGGGCTCAGAAATGGAGCTCCTGGGCTGTGAAGTGAAGCCC	1762
Db	2296	GGTGGTGGCCAGCAGGCATCGGGATGAGGCTGGAGGTGCTGGGCTGTGACTGGGACAGACT	2355
Qy	1763	CTA 1765	
Db	2356	CAA 2358	

RESULT 14

US-08-936-135-23
; Sequence 23, Application US/08936135
; Patent No. 6054293
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: He, Zhigang
; APPLICANT: Chen, Hang
; TITLE OF INVENTION: Semaphorin Receptors
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS: 26
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,135
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC97-288-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4780 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-296-264-33-1
Query Match 19.4%; Score 537.8; DB 3; Length 4780;
Best Local Similarity 58.9%; Pred. No. 1.7e-151;
Matches 992; Conservative 0; Mismatches 667; Indels 24; Gaps 3;
QY 107 CCGGGTACCTTACATCTCCCTGGTATCTCTCATCTTATCATCCCAAGTGAATAATGCGAAT 166
Db 676 CTGGCTACATCACTTCCCGAGCTACCCCGAGGACTATCCCTCCACCAAGACTGTGAGT 735
QY 167 GCCTGATTCAGCTCCGGACCCATACAGAGAATATGATCACTTCAACCCCTCACTTCG 226
Db 736 GGATTGCTACGCCCGGACCCCAACCAAGAGATTTGTTCTCAACTTCAACCCCTCACTTTG 795
QY 227 ATTTGGAGGACAGACTGCAAGTATGACTACGTGGAAGTCTTTCGATGGAGAAATGAAA 286
Db 796 AATCGAGAACAGGACTGCAAGTATGACTTCAATGATTCGGGATGGGACAGTGAAT 855
QY 287 ATGGACATTTTGGGAAAGTTCTGCGAAAGATAGCCCTCTCTCTGTTGTTCTTCAG 346
Db 856 CAGCTGACCTCTCTGGGCAAGCACTGTGGGAAACATCGCCCGCCACCATCTCTCTCAG 915
QY 347 GGCATTTCTTTTATCAAAATTTGCTCTGACTACGAACACAGATGTGACGAGATTTTCCA 406
Db 916 GCTCGGTGTTATACATCAAGTTCACTTCACTTACGCTACGCCCGGAGGGGAGGTTTCTCTC 975
QY 407 TAGCTTATGAAATTTTCAAGAGAGGTCCTGAA---TGTTCCCAAGAACTACACAAACCTTA 463
Db 976 TAGCTTATGAGATCTTCAAAACAGGCTCTGAAGATTTTCCAGAACTTTACAAAGCCCA 1035
QY 464 GTGGATGATAAGTCCCGGATTCCTCGTGAATAATATCCCAACAGCTTGAATGACATT 523
Db 1036 ATGGGACCATTTGAATCTCCAGGTTTCCAGAGAGTATCCACAAATCTGGAGTGTACCT 1095
QY 524 ATATTGTTCTTGGCCAAAGATCTCAGAGATTTATCTGGAATTTGAAGCTTTGACCTGG 583
Db 1096 TCACATCTGGCCAAACCCAGGATGGAGATCATCTCAGATTCTCGACTTTGACCTGG 1155
QY 584 AGCTGACTCAAAATCTCCAGGGGGATGTCTGTGCTTACGACCGGCTAGAAATCTGGG 643
Db 1156 AGCATGACCTCTACAAGTGGGGAGGAGAGACTGTAAATATGACTGGCTGGACATCTGG 1215
QY 644 ATGATTCCTGTATGTTGCCCTTACATTTGGGGGTTTACTGTGGACAGAAACACCAAGTTC 703
Db 1216 ATGGCATTCACATGTTGGACCTCTGATTGGCAAGTACTGTGGGACGAAACACCCCTCCA 1275
QY 704 GAATCGATCTCATCGGCATTTCTCTCCATGTTTTCACACCGACAGCGGATAGCAA 763
Db 1276 AACTCCGCTGCTCCAGGGATCTCTCTCTGACCTTTTACACGGAGACATGGCAGTGGCCA 1335
QY 764 AAGAAGTTTCTCAGCAAACTACAGTGTCTTGCAGAGCAGTGTCTCAGAAAGATTTCAAAAT 823
Db 1336 AGGATGGCTTCTCCGACAGTACTATTGATCCACAGGACCACTGTAGATTTTCACT 1395
QY 824 GTATGAAGCTCTGGCATGGATCAGGAGAAATTCATTCTGACAGATCACAGTTCCT 883
Db 1396 GCAATGCTCCTTTGGGAATGGAGTCTGGCCGGATGCTTAATGAACAGATCAGTGTCTCT 1455
QY 884 ---CCAGTATAGCAACCACTGGTCTGACAGCGCTCCCGCTGAACTACCTCTGAGATG 940
```

```
Db 1456 CCACCTTCTCTGATGGGAGGTGGACTCTCTCAACAGAGCCGCTCCATGTTGATGACAATG 1515
QY 941 GGTGGACTCCCGAGAGGATTTCTTACCGAGAGTGGATACAGGTAGACTTGGGCTTCTGC 1000
Db 1516 GCTGGACACCAATTTGGATTTCAACAGAGGATATCTCCAGGTGGACTCGCTTCTCTAA 1575
QY 1001 GCTTTGTACGGCTGTGCGGACACAGGCGCCCATTTTCAAAAGAAACCAAGAAATATT 1060
Db 1576 CCATGCTCAGACCAATGCAACACAGGAGGCCATTTCCAGGGAACCCAGAGGCTACT 1635
QY 1061 ATGTCAAGACTTACAAGATCGAGTGTAGTCTCAACAGGGGAGAGACTGATCACATATAAG 1120
Db 1636 ACGTCAAAATCGTACAGCTGGAAGTACGACAAATGGTGAAGATTGGATGTTCTACCGGC 1695
QY 1121 AAGGAACAAACCTGTTCTCTTTTCAGGGAACACCAACCCACAGATGTTGTGTTGTCAG 1180
Db 1696 ATGGCAAAACCAACACAGATATTTCAAGCGAAACAATGATCGACCGAGGTGGTCTAAACA 1755
QY 1181 TATTCCCAAAACCACTGATAACTCGATTTTGTCCGAATCAAGCCTGCAACTTGGGAAATG 1240
Db 1756 AGCTCCACATGCCACTGCTGACTCGGTTTATCAGGATCCGCCCGCAGACGTTGSCATTGG 1815
QY 1241 GCATATCTATGAGATTTGAAGTATACGGTTGCAAGATAACAGATATFCTTGTGCTCTGAA 1300
Db 1816 GCATTTGCCCTTCCGCTGGAGCTCTTTGGCTCGCGGTTACAGATGACACCTGCTCCAACA 1875
QY 1301 TGTGTTGGTATGGTGTCTGGACTTATTTCTGACTCCAGATCCACATCATCATCAACCAAGGAG 1360
Db 1876 TGTCTGGGATGCTCTCGGCTCATTTGCTGTATACCCAGATCTCTGCTCTCTCCACCCGAG 1935
QY 1361 ACAGAAACTGGATGCTGTAACATCCGCTGGTGTACCACTGCTCTGCTGGCTGGGCACTTC 1420
Db 1936 AGTACCTCTGGAGCCCGCAGTGTGCTGCCGCTGTTAGTAGCCGCTCTGCTGGTGTTCCTC 1995
QY 1421 CACCCGCACTCATCTCTACATCAATGAGTGGCTCCAAATAGACCTTGGGGAGGAGAAGA 1480
Db 1996 GGAACCTCTCAAGCCAGCCAGAGTGAAGATGGCTTCAGGTAGACCTGGGACACCCCAAGA 2055
QY 1481 TCGTGGGGGCACTCATCATTCAGGTGGGAAGCACCAGAGAGAACA----- 1526
Db 2056 CAGTGAAGGGGTCACTATCCAGGGAGCCGAGGAGGACAGCATCATCTGCGTGGAG 2115
QY 1527 ---GGTGTTCATGAGGAAGTTCAAGATCGGTGACAGCAACACAGGCTCGGACTGGAAGA 1582
Db 2116 CCAGGGCGTTGTACCAAGTTCAAAAGTCTCTTACAGCTTAAATGGCAAGGACTGGGAAT 2175
QY 1583 TGATCATGATGACACCAACCAAGCCGAAGTCTTTTGGGGCAACACACATATGATA 1642
Db 2176 ATATCCAGGACCCAGGACTCAGACAAAGCTGTTTGAAGGGAACATGCACTATGACA 2235
QY 1643 CACCTGAGCTGGGACTTTTTCAGCTCTCTCCAGCGATTTCATCAGGATCTACCCCGAGA 1702
Db 2236 CCCCCTGACATCCGAGGTTGATCTCTGTTCCAGCGGAGTATGTGCGGGTGTACCCAGAGA 2295
QY 1703 GAGCCACTCATCGGAGCTGGGGCTCAGAATGAGCTGCTGGGCTGTGAAGTGAAGCC 1762
Db 2296 GGTGTGCGCAGCAGGCACTCGGATGAGGCTGGAGTGTGGCTGTGACTGTGACAGACT 2355
QY 1763 CTA 1765
Db 2356 CAA 2358
```

RESULT 15

```
US-08-480-229C-9
; Sequence 9, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quermous, Thomas
; APPLICANT: Hogan, Bridg
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY - REGULATED ENDOTHELIAL
```

Query Match Best Local Similarity Matches	2.5%; Score 69.6; DB 2; Length 2303; 45.2%; Pred. No. 1.4e-10; 396; Conservative	0; Mismatches 459; Indels 21; Gaps
QY	924	GAACTACCCCTCAGAAATGGGTGGACTCCCGGAGAGAGATTCTTACCAGAGTGGATACAGGT 983
DB	1212	GAAGGCCCTTATAAATGCCCTGGACAGCTGCTGCAAAATGCACAGATGGCCATGGATTTCAGAT 1271
QY	984	AGACTTGGGCGCTTGCGCTTTGTACGGCTGTGCGGACACAGGCGGCCATTTTCAAAGA 1043
DB	1272	AATTTCCAAGAAAAATGAGAGHCACTGGTGTATTATCCCAGAGGACAAAAGGATTGG 1331
QY	1044	AACCAGAAGAAATATTATGTCAAGACTTTACAAGATCGACCTTAGCTTCAACGGGGAGA 1103
DB	1332	AAGCCAGATACATAAAATCCTACAAATTTGCCCTACAGCAATGCGGGAAGACCTGGGC 1391
QY	1104	CTGGATCACCATAAAGAAGAAACAACCTGTTCTCTTTCAGGGAACAACCAACCCAC 1163
DB	1392	AATGTACAAAGTAAGAGCCACCAATGAAGAGATGGTCTTCTCGTGGAAATGTTGATAACA 1451
QY	1164	AGATGTTGTGGTGGCAGTATTTCCCCAACCACTGATTAACCTCGATTGTCTCCGAATCAAGCC 1223
DB	1452	CACACATATGCTAAATTCTTTTCACACCCCCCAATCAAAGCTCAGTATGTGAAGACTCTACCC 1511
QY	1224	TGCAACTTGGGAACTGGCATATCTATGAGATTTGAAAGTATACGGTTTCCAAGATAACAGA 1283
DB	1512	CCAAATTTGTGCAAGGCATGTGACTTTAAGNAATGGAACCTCTTGGCTGTGAGCTCTCAGG 1571
QY	1284	TTATPCTTGTCTTGGAAATGTTGGGTATGGTGTCTGGACTTTATTTCTGACTCCCGAGATCAC 1343
DB	1572	CTGT-----TCAGAACCTTTGGGGATCAAAATCAGGGCATATACAAGACTACCAGATCAC 1625
QY	1344	ATCATCCAACCAAGGACAGAAACTGGATG~-CCTGAAAACATCCCGCTTGGTAAACAGT 1401
DB	1626	TGCCTCAGGGCTCTTCAGAACACTTCAACATGGACATGTTTTACTTTGGGAACCAAGGAAAGC 1685

THIS PAGE BLANK (USPTO)

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2762.4	99.7	5653	12	US-10-104-440-1		Sequence 1, Appl1
2	549.6	19.8	3404	12	US-10-104-440-3		Sequence 3, Appl1
3	306.4	11.1	430	10	US-09-864-761-1508		Sequence 1508, Ap
4	279.4	10.1	319	10	US-09-864-761-20513		Sequence 20513, A
5	247.4	8.9	422	10	US-09-864-761-3746		Sequence 3746, Ap
6	237.8	8.6	241	10	US-09-864-761-21153		Sequence 21153, A
7	215.2	7.8	246	10	US-09-867-701-3388		Sequence 3388, Ap
8	195	7.0	195	10	US-09-864-761-18267		Sequence 18267, A
9	188.4	6.8	190	10	US-09-864-761-20524		Sequence 20524, A
10	176.2	6.4	397	10	US-09-864-761-4405		Sequence 4405, Ap
11	169.2	6.1	389	10	US-09-864-761-3757		Sequence 3757, Ap
c 12	88.4	3.2	487	10	US-09-864-761-2952		Sequence 2952, Ap
c 13	87	3.1	251	10	US-09-864-761-19733		Sequence 19733, A
c 14	80.4	2.9	482	10	US-09-864-761-115		Sequence 115, App
c 15	78.8	2.8	241	10	US-09-864-761-16948		Sequence 16948, A
16	62.2	2.2	1270	10	US-09-880-107-3020		Sequence 3020, Ap
17	56.2	2.0	6909	10	US-09-880-107-2275		Sequence 2275, Ap
18	55	2.0	4771	10	US-09-942-366-2		Sequence 2, Appl1
19	51.8	1.9	3919	10	US-09-942-366-4		Sequence 4, Appl1

|||||
Db 359 TCTCCCTGGTATCTCTCATTTATATACCAAGTGAAAAATGCGAATSGCGTATTCAGGCT 418
QY 181 CCGGACCCATACAGAGAATATGATCAACTTCAACCCCTCACCTTCGATTTGGAGGACAGA 240
Db 419 CCGGACCCATACAGAGAATATGATCAACTTCAACCCCTCACCTTCGATTTGGAGGACAGA 478
QY 241 GACTGCAAGTATGACTACGTGGAGCTTTCGATGGAGAAAAATGAAAAATGGACATTTTAGG 300
Db 479 GACTGCAAGTATGACTACGTGGAGCTTTCGATGGAGAAAAATGAAAAATGGACATTTTAGG 538
QY 301 GGAAGTTCTGTGGAAAGATAGCCCTCTCTCTGTTGTCTTCAGGGCCATTTCTTTT 360
Db 539 GGAAGTTCTGTGGAAAGATAGCCCTCTCTCTGTTGTCTTCAGGGCCATTTCTTTT 598
QY 361 ATCAAAATTTGTCCTGACTACGAAACACATGGTGCAGGATTTTCCATAGCTTATGAAAT 420
Db 599 ATCAAAATTTGTCCTGACTACGAAACACATGGTGCAGGATTTTCCATAGCTTATGAAAT 558
QY 421 TTCAGAGAGGTCCTGAATGTTTCCAGAACTACACAAACACCTAGTGGAGTGATAAAGTCC 480
Db 659 TTCAGAGAGGTCCTGAATGTTTCCAGAACTACACAAACACCTAGTGGAGTGATAAAGTCC 718
QY 481 CCGGGATTCCTTGAAAAATATCCCAACAGCCTTGAATGCACCTTATATGTTCTTTGGGCCA 540
Db 719 CCGGGATTCCTTGAAAAATATCCCAACAGCCTTGAATGCACCTTATATGTTCTTTGGGCCA 778
QY 541 AAGATGTCAGAGATTTATCTGGAAATTTGAAAGCTTTGAGCTTGAGGCTGACCTCAAAATCCT 600
Db 779 AAGATGTCAGAGATTTATCTGGAAATTTGAAAGCTTTGAGCTTGAGGCTGACCTCAAAATCCT 838
QY 601 CCAGGGGGATGTTCTGCTACGACCGCTAGAAATCTGGATGGATTCCTGATGTT 660
Db 839 CCAGGGGGATGTTCTGCTACGACCGCTAGAAATCTGGATGGATTCCTGATGTT 898
QY 661 GGCCTTCATTTGGGGCTTACTGTGGACAGAAACACAGGTCGAAATCCGATCCTCATCG 720
Db 899 GGCCTTCATTTGGGGCTTACTGTGGACAGAAACACAGGTCGAAATCCGATCCTCATCG 958
QY 721 GGCAATCTCTCCATGTTTATACCGGACAGCGGATAGCAAAAGAGTTTCTCAGCA 780
Db 959 GGCAATCTCTCCATGTTTATACCGGACAGCGGATAGCAAAAGAGTTTCTCAGCA 1018
QY 781 AACTACAGTGTCTGCAGAGCAGTGTCTCAGAAATTTTCAAAATGTATGGAAGCTCTGGGC 840
Db 1019 AACTACAGTGTCTGCAGAGCAGTGTCTCAGAAATTTTCAAAATGTATGGAAGCTCTGGGC 1078
QY 841 ATGGAATCAGGAGAAATTCATTTGACCCAGATCACAGCTTCTTCCAGTATAGCCCAAC 900
Db 1079 ATGGAATCAGGAGAAATTCATTTGACCCAGATCACAGCTTCTTCCAGTATAGCCCAAC 1138
QY 901 TGGTCTGCAGAGCGCTCCCGCTGAACTACCTGAGAATGGGTGGACTCCCGGAGAGGAT 960
Db 1139 TGGTCTGCAGAGCGCTCCCGCTGAACTACCTGAGAATGGGTGGACTCCCGGAGAGGAT 1198
QY 961 TCTTACCGAGATGGATACAGGTAGACTTTGGCCCTTCTCGCTTTGTTCACGGCTGTGCGG 1020
Db 1199 TCTTACCGAGATGGATACAGGTAGACTTTGGCCCTTCTCGCTTTGTTCACGGCTGTGCGG 1258
QY 1021 ACACAGGGCGCATTTTCAAAAGAAACCAAGAAATATATGTCAAGACTTTACAAGATC 1080
Db 1259 ACACAGGGCGCATTTTCAAAAGAAACCAAGAAATATATGTCAAGACTTTACAAGATC 1318
QY 1081 GAGCTTAGTCTCAAGGGGAAGACTGGATACCATTAAGAAAGGAAACAAACCTGTTCTC 1140
Db 1319 GAGCTTAGTCTCAAGGGGAAGACTGGATACCATTAAGAAAGGAAACAAACCTGTTCTC 1378
QY 1141 TTTTACGGGAAACCAACCCACAGATGTTGGTTGAGTATTCCTCCAAACCACTGATA 1200
Db 1379 TTTTACGGGAAACCAACCCACAGATGTTGGTTGAGTATTCCTCCAAACCACTGATA 1438
QY 1201 ACTCGATTTGCCGAATCAGGCTCGAATCTGGGAACTGGGCATATCTATGAGATTTGAA 1260
|||||

Db 1439 ACTCGATTTGTCCGAATCAAGCCTGCAACTTGGGAAACTGGCATATCTATGAGATTTGAA 1498
QY 1261 GTATACGGTTGCAAGATAACAGATTATCTGTCTCTGGAAATGTTGGTATGTTGTTCTGGA 1320
Db 1499 GTATACGGTTGCAAGATAACAGATTATCTGTCTCTGGAAATGTTGGTATGTTGTTCTGGA 1558
QY 1321 CTTATTTCTGACTCCCATGATCATCATCAACCAAGGAGACAGAAACTGGATGCTGAA 1380
Db 1559 CTTATTTCTGACTCCCATGATCATCATCAACCAAGGAGACAGAAACTGGATGCTGAA 1618
QY 1381 AACATCCGCTGGTAACCACTGCTCTGCTGGGACTTCCACCCGACCTCATTTCTCTAC 1440
Db 1619 AACATCCGCTGGTAACCACTGCTCTGCTGGGACTTCCACCCGACCTCATTTCTCTAC 1678
QY 1441 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAACTCGTAGGGGCAATCATCAT 1500
Db 1679 ATCAATGAGTGGCTCCAAATAGACCTGGGGAGGAGAACTCGTAGGGGCAATCATCAT 1738
QY 1501 CAGGTTGGGAAGCACCGAGAGAACTGTTTCATGAGAACTTCAAGATCGGGTACAGC 1560
Db 1739 CAGGTTGGGAAGCACCGAGAGAACTGTTTCATGAGAACTTCAAGATCGGGTACAGC 1798
QY 1561 AACACGGCTCGGACTGGGAAGATGATCATGATGACAGCAACGCAAGCGGAAGTCTTTT 1620
Db 1799 AACACGGCTCGGACTGGGAAGATGATCATGATGACAGCAACGCAAGCGGAAGTCTTTT 1858
QY 1621 GAGGCAACAACTATGATACACTGAGCTGCGGACTTTTCCAGCTCTCTCCACGCA 1680
Db 1859 GAGGCAACAACTATGATACACTGAGCTGCGGACTTTTCCAGCTCTCTCCACGCA 1918
QY 1681 TTCATCAGGATCTACCCGAGAGACCACTCATGCGGACTGGGCTCAGAAATGGAGCTG 1740
Db 1919 TTCATCAGGATCTACCCGAGAGACCACTCATGCGGACTGGGCTCAGAAATGGAGCTG 1978
QY 1741 CTGGGCTGTGAAGTGGAGCCCTTACAGCTGGACCCACCTCCCAACGGGAACTTGGTG 1800
Db 1979 CTGGGCTGTGAAGTGGAGCCCTTACAGCTGGACCCACCTCCCAACGGGAACTTGGTG 2038
QY 1801 GATGAATGTGATGAGCAGCCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 1860
Db 2039 GATGAATGTGATGAGCAGCCCAACTGCCACAGTGGAAACAGGTGATGACTTCCAGCTC 2098
QY 1861 ACAGTGGCACCACTGTGCTGGCCACAGAAAAACCCAGGTCATAGACAGCACCATACAA 1920
Db 2099 ACAGTGGCACCACTGTGCTGGCCACAGAAAAACCCAGGTCATAGACAGCACCATACAA 2158
QY 1921 TCAGAGTTTCCAAACATATGGTTTTAACTGTGAATTTGGCTGGGGCTCTCACAGACCTTC 1980
Db 2159 TCAGAGTTTCCAAACATATGGTTTTAACTGTGAATTTGGCTGGGGCTCTCACAGACCTTC 2218
QY 1981 TGCCACTGGGAACATGACAATCACCTGAGCTCAAGTGGAGTGTGTTGACCAAGCAAGCG 2040
Db 2219 TGCCACTGGGAACATGACAATCACCTGAGCTCAAGTGGAGTGTGTTGACCAAGCAAGCG 2278
QY 2041 GGACCCATTCAGGATCACACAGGAGATGSCAACTTCATCTATTTCCCAAGCTGACCAAAAT 2100
Db 2279 GGACCCATTCAGGATCACACAGGAGATGSCAACTTCATCTATTTCCCAAGCTGACCAAAAT 2338
QY 2101 CAGAAGGGCAAGTGGCTGCGCTGGTGGAGCTGTGGTTTTATTTCCAGAACTCTGCCAC 2160
Db 2339 CAGAAGGGCAAGTGGCTGCGCTGGTGGAGCTGTGGTTTTATTTCCAGAACTCTGCCAC 2398
QY 2161 TGCATGACCTTCTGCTATCACATGCTGGTCCCGCTGGGACACTCAGGGTCAAACTG 2220
Db 2399 TGCATGACCTTCTGCTATCACATGCTGGTCCCGCTGGGACACTCAGGGTCAAACTG 2458
QY 2221 CGCTACCAAGGCCAGGAGTACGATCAGCTGGTCTGGATGGCCATTTGGACACCAAGGT 2280
Db 2459 CGCTACCAAGGCCAGGAGTACGATCAGCTGGTCTGGATGGCCATTTGGACACCAAGGT 2518
QY 2281 GACCACTGGAAGGAGGGGTGCTTTCCTCCAAAGTCTCTGAAACTTTATCAGGTGATT 2340
Db 2519 GACCACTGGAAGGAGGGGTGCTTTCCTCCAAAGTCTCTGAAACTTTATCAGGTGATT 2578
|||||

Qy	2341	TTCCGAGGGCGGAATTCGGAAAGGAAACCTTGGTGGGATTGCTGTGGATGACATTAGTATT	2400
Db	2579	TTCCGAGGGCGGAATTCGGAAAGGAAACCTTGGTGGGATTGCTGTGGATGACATTAGTATT	2638
Qy	2401	AATTAACCCACATTTTACACAAGAATTTGCAAAACACAGACAGACCTCGGATTAATAAACAACCCA	2460
Db	2639	AACAACCCACATTTTACACAGAAGATTTGCAAAACACAGACAGACCTCGGATTAATAAACAACCCA	2598
Qy	2461	GAATTTAAATTTGATGAACACAGGGAGCACGCCAGGATACGAAGGTTGAAGGAGAAGGTCAC	2520
Db	2699	GAATTTAAATTTGATGAACACAGGGAGCACGCCAGGATACGAAGGTTGAAGGAGAAGGTCAC	2758
Qy	2521	AAGAACATCTCCAGGNAAGCCAGGCCAATGTGTTGAAGACCTTTAGAAACCATCCTCATCACC	2580
Db	2759	AAGAACATCTCCAGGNAAGCCAGGCCAATGTGTTGAAGACCTTTAGATCCCATCCTCATCACC	2818
Qy	2581	ATCATAGCCATGAGCCCTCTGGGGTCTCCTCTGGGGGCTGCTGTGGGGTCTGCTGCTGAC	2640
Db	2819	ATCATAGCCATGAGTCCCTCTGGGGTCTCCTCTGGGGGCTGCTGTGGGGTCTGCTGCTGAC	2878
Qy	2641	TGTGCTCTGTGGCATTAATGGGATGTACAGAAAGAACTTTGCTGCCCTGGAGAACTATAAC	2700
Db	2879	TGTGCTCTGTGGCATTAATGGGATGTACAGAAAGAACTTTGCTGCCCTGGAGAACTATAAC	2938
Qy	2701	TTTGAACCTTGTGGATGCTGTAAGTTGAAAAAGACAAACTGAAATACACAGAGTACTTAT	2760
Db	2939	TTTGAACCTTGTGGATGCTGTAAGTTGAAAAAGACAAACTGAAATACACAGAGTACTTAT	2998
Qy	2761	TCGGAGGCGATGA 2772	
Db	2999	TCGGAGGCGATGA 3010	

RESULT 2

```

US-10-104-440-3
; Sequence 3, Application US/1010440
; Patent No. US20020132774A1
; GENERAL INFORMATION:
; APPLICANT: KLAGSBRUN, Michael
; APPLICANT: SOKER, Shay
; APPLICANT: MIAO, Hua Quan
; TITLE OF INVENTION: ANTAGONISTS OF NEUROFILIN RECEPTOR FUNCTION AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 48802 C
; CURRENT APPLICATION NUMBER: US/10/104,440
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/580,803
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/069,155
; PRIOR FILING DATE: 1997-12-09
; PRIOR APPLICATION NUMBER: 60/069,687
; PRIOR FILING DATE: 1997-12-29
; PRIOR APPLICATION NUMBER: 60/078,541
; PRIOR FILING DATE: 1998-03-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 3404
; TYPE: DNA
; ORGANISM: human
US-10-104-440-3

```

Query Match	19.8%	Score 549.6;	DB 12;	Length 3404;
Best Local Similarity	54.0%;	Pred. No. 3.2e-163;		
Matches 1498; Conservative	0;	Mismatches 1139;	Indels 135;	Gaps 13;

[illegible]

Db	238	GACTACCCCTCCACACAGAACTCGCGAGTCGGATTGTTTACGCCGCCCAACCCCAACCAAG	297
Qy	199	ATTATGATCAACTTCAACCCCTCACTTTCGATTTGGGAGCAGAGACTGCAAGTATGACTAC	258
Db	298	ATTGCTCTCAACTTCAACCCCTCACTTGGAAATCGAGAGCAGCAGTGCAGTATGACTTT	357
Qy	259	GTGGAAGTCTTCGATCGAGAGAAATGAAATGGACATTTTAGGGGAAAGTTCCTGTGGAAG	318
Db	358	ATCGAGATTCCGGATGGGACAGTAGAATCCGCAGACCTCCTGGGCAAAACACTGTGGGAAC	417
Qy	319	ATAGCCCTCCTCCTGTGTGCTTCAGGGCCATTTCTTTTATCAAAATTTGTCTCTGAC	378
Db	418	ATGCCCGCCGCCACCATCATCTCTCGGGCTCCATGCTCTACATCAAGTTCACCTCCGAC	477
Qy	379	TACGAAACATAGTGCAGGATTTTCCATACGTTATGAAATTTTCAAGAGAGGCTCTCAA	438
Db	478	TACGCCCGCAGGGCAGGCTTCTCTCTCGCTACGAGATCTTCAAGACAGGCTCTGAA	537
Qy	439	---TGTTCCCAGAACTACACACACCTAGTGGAGTGATAAAGTCCCCCGGATTCCTCTGAA	495
Db	538	GATTGTCTCAAAAACCTTCAACAGCCCAACGGGACCATCGAATCTCTCTGGTTCCTCTGAG	597
Qy	496	AAATATCCCAACAGCTTGAATGCACCTATATTGTCTTTGGGCCAAAGATGTCAGAGATT	555
Db	598	AAGTATCCACACAACTTGGACTGCACCTTTTACATCTCGGCCAAACCCAAAGATGAGATC	657
Qy	556	ATCCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAAATCTCTCCAGGGGGATGTC	615
Db	658	ATCCTGCAGTTCCTGATCTTTGACCTGGAGCATGACCTTTCCAGTGGGAGAGGGGAC	717
Qy	616	TGTCGCTACGACCGGCTAGAAATCTGTGGGATGGAATCCCTGATGTGTGGCCCTCACATTTGGG	675
Db	718	TGCAAGTACGATTGGCTGGACATCTGGGATGGCATTCACATGTGTGGCCCCCTGATTGGC	777
Qy	676	CGTTACTGTGCGACAGAAACACACAGTCGAATCCGATCCTCATCGGGATTTCTCCCATG	735
Db	778	AAGTACTGTGGGACCAAAACCCCTGAACTTCTGATTCATCGAGGGGATCCTCTCCCTG	837
Qy	736	GTTTTTTACACCGACGCGGATAGCAAAAGAGTGTCTCAGCAAACTACAGTGTCTTG	795
Db	838	ACCTTTCACAGCAGCATGGGGTGGCCAGGATGGCTTCTCTGCGCGTTACTACCTGGTC	897
Qy	796	CAGAGCAGTGTCTCAGAGAATTTCAAATGTATGGAAGTCTGGGATCGGAATCAGGAGAA	855
Db	898	CACCAAGAGCCACTAGAGAACTTTCAGTGCAATGTTCCTCTGGGCATGGAGCTCTGGCCG	957
Qy	856	ATTCTATCTACAGAGATCAGAGCTTCTT---CCCAGTATAGCACCAACTGGTCTGCGAG	912
Db	958	ATTGCTAATGAACAGATCAGTGCCCTCATCTAGCTACTGTATGGGAGGTGGACCCCTCAA	1017
Qy	913	CGCTCCCGCTGAACTACCTCAGAAATGGTGGACTCCCGGAGAGGATTTCTTACCGAGAG	972
Db	1018	CAAGCGGCTCCATGGTGATGACATGGCTGGACCCCACTTTGGATTTCCAAACAGAG	1077
Qy	973	TGGATACAGATAGACTTGGGCCCTTCTGCGCTTTGTACGCGCTGTGGGACACAGGGGCC	1032
Db	1078	TATCTCCAGTGGACCTGGCGCTTTTAAACATGCTCACGGCCATCGCAACACAGGGAGCG	1137
Qy	1033	ATTTCAAAAGAAACCAAGAGAATATTATGTCAAGACTTACAGATCGAGCTTAGCTTAC	1092
Db	1138	ATTTCCAGGAAACACAGAAATGGCTACTACGTCAAACTCTACAGCTGGAAGTCAAGACT	1197
Qy	1093	AACGGGAAGACTGGATCACCATAAAGAAAGAAACAACTGTCTCTTTTCAGGGAAAC	1152
Db	1198	AATGGAGAGGACTGGATGGTGTAACGGCATGCAAAACCAAGGATATTTCAAGCCAC	1257
Qy	1153	ACCAACCCCAAGATGTTGTGGTTCAGTATTTCCCCAAACCACTGATAACTCGATTGTC	1212
Db	1258	AACGATGCACTGAGTGGTCTTGAACAGCTCCAGCTCCACTGCTGACAAAGTTGTT	1317
Qy	1213	CGAATCAAGCCTCGAATCTGGGAACTGGCATATCTATGAGATTTGAAGTATACGGTTGC	1272
Db	1318	AGAAATCCGCTCAGACCTTGGCACTCAGGATATGCGCTTCCGCTCGAGCTCTTCGGCTGC	1377

;; PRIORITY APPLICATION NUMBER: PCT/US01/00663
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00662
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00661
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00670
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: US 60/234,687
;; PRIORITY FILING DATE: 2000-09-21
;; PRIORITY APPLICATION NUMBER: US 09/608,408
;; PRIORITY FILING DATE: 2000-06-30
;; PRIORITY APPLICATION NUMBER: US 09/774,203
;; PRIORITY FILING DATE: 2001-01-29
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 1508
;; LENGTH: 430
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL121748.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.8
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.7
US-09-864-761-1508

Query Match 11.1%; Score 306.4; DB 10; Length 430;
Best Local Similarity 99.7%; Pred. No. 8.8e-87;
Matches 307; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1280 CAGATTATCTGCTCGAATGTTGGGTATGTTGGTCTGGACTTATTCTGACTCCAG 1339
|||||
DB 123 CAGATTATCTGCTCGAATGTTGGGTATGTTGGTCTGGACTTATTCTGACTCCAG 182
|||||
QY 1340 TCACATCATCCAAACCAAGGAGACAGAACTGGATGCTGAAACATCCCGCTGGTAACCA 1399
|||||
DB 183 TCACATCATCCAAACCAAGGAGACAGAACTGGATGCTGAAACATCCCGCTGGTAACCA 242
|||||
QY 1400 GTGCTCTGCTGGGACATTCACCCGACCTCATCTTCTACATCATGATGCTCCAA 1459
|||||
DB 243 GTGCTCTGCTGGGACATTCACCCGACCTCATCTTCTACATCATGATGCTCCAA 302
|||||
QY 1460 TAGACCTGGGGAGGAGAGATCGTGGGGGATCATCTTACAGGTGGGAAGCACCAG 1519
|||||
DB 303 TAGACCTGGGGAGGAGAGATCGTGGGGGATCATCTTACAGGTGGGAAGCACCAG 362
|||||
QY 1520 AGAACAGGTGTTTCATGAGAGATTCAGATCGGGTACAGCAACACGCTCGGACTGGA 1579
|||||
DB 363 AGAACAGGTGTTTCATGAGAGATTCAGATCGGGTACAGCAACACGCTCGGACTGGA 422
|||||
QY 1580 AGATGATC 1587
|||||
DB 423 AGATGATC 430

RESULT 4
US-09-864-761-20513
; Sequence 20513, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

;; FILE REFERENCE: Aeonica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIORITY APPLICATION NUMBER: US 60/180,312
;; PRIORITY FILING DATE: 2000-02-04
;; PRIORITY APPLICATION NUMBER: US 60/207,456
;; PRIORITY FILING DATE: 2000-05-26
;; PRIORITY APPLICATION NUMBER: US 09/632,366
;; PRIORITY FILING DATE: 2000-08-03
;; PRIORITY APPLICATION NUMBER: GB 24263.6
;; PRIORITY FILING DATE: 2000-10-04
;; PRIORITY APPLICATION NUMBER: US 60/236,359
;; PRIORITY FILING DATE: 2000-09-27
;; PRIORITY APPLICATION NUMBER: PCT/US01/00666
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00667
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00664
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00669
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00665
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00668
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00663
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00662
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00661
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: PCT/US01/00670
;; PRIORITY FILING DATE: 2001-01-30
;; PRIORITY APPLICATION NUMBER: US 60/234,687
;; PRIORITY FILING DATE: 2000-09-21
;; PRIORITY APPLICATION NUMBER: US 09/608,408
;; PRIORITY FILING DATE: 2000-06-30
;; PRIORITY APPLICATION NUMBER: US 09/774,203
;; PRIORITY FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 20513
;; LENGTH: 319
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL121748.2
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
;; OTHER INFORMATION: EST_HUMAN HIT: AA987327.1, EVALUATE 0.00e+00
;; OTHER INFORMATION: SWISSPROT HIT: O14786, EVALUATE 1.00e-51
;; OTHER INFORMATION: NT HIT: AF280547.1, EVALUATE 0.00e+00
US-09-864-761-20513

Query Match 10.1%; Score 279.4; DB 10; Length 319;
Best Local Similarity 99.6%; Pred. No. 2.5e-78;
Matches 280; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1334 CCCAGATCACATCATCCAAACCAAGGAGAGAGAACTGGATGCTGAAACATCCCGCTGG 1393
|||||
DB 1 CCCAGATCACATCATCCAAACCAAGGAGAGAGAACTGGATGCTGAAACATCCCGCTGG 60
|||||
QY 1394 TAACCATGCTGCTGGCTGGGACCTTCCACCCGACCTCATCTTCTACATCAATGAGTGGC 1453
|||||
DB 61 TAACCATGCTGCTGGCTGGGACCTTCCACCCGACCTCATCTTCTACATCAATGAGTGGC 120
|||||
QY 1454 TCCAAATAGACCTGGGGAGGAGAGAGATCGTGAGGGGATCATCTTACAGGTGGGAAGC 1513
|||||
DB 121 TCCAAATAGACCTGGGGAGGAGAGAGATCGTGAGGGGATCATCTTACAGGTGGGAAGC 180
|||||
QY 1514 ACCGAGAGAAAGGTGTTTCATGAGGAAGTTCAAGATCGGGTACAGCAACACGCTCGG 1573

```
Db 181 ACCGAGAGAACAGGTGTTTCATGAGGAGTTCAGATCGGTCAGCAACACGGCTCGG 240
QY 1574 ACTGGAAGATGATCATGATGACAGCAACGCAAGCGGAAG 1614
Db 241 ACTGGAAGATGATCATGATGACAGCAACGCAAGCGGAAG 281

RESULT 5
US-09-864-761-3746
; Sequence 3746, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3746
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121748.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
US-09-864-761-3746
Query Match 8.9%; Score 247.4; DB 10; Length 422;
```

```
Best Local Similarity 99.6%; Pred. No. 4.6e-68;
Matches 248; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1280 CAGATTATCTCTGCTCTGGAATGTTGGGTATGTTGCTGACTTATTCTGACTCCCGA 1339
Db 174 CAGATTATCTCTGCTCTGGAATGTTGGGTATGTTGCTGACTTATTCTGACTCCCGA 233

QY 1340 TCACATCATCCAAACGAGAGACAGAACTGGATGCTGAAAAACATCCGCTGGTAACCA 1399
Db 234 TCACATCATCCAAACGAGAGACAGAACTGGATGCTGAAAAACATCCGCTGGTAACCA 293

QY 1400 GTCGCTCTGGCTGGGCACATTCACCCGCACCTCATTCCTACATCAATGAGTGGCTCAA 1459
Db 294 GTCGCTCTGGCTGGGCACATTCACCCGCACCTCATTCCTACATCAATGAGTGGCTCAA 353

QY 1460 TAGACCTGGGGGAGGAGAGATCGTGAGGGGCATCATCTTCAGGTGGGAACACCGAG 1519
Db 354 TAGACCTGGGGGAGGAGAGATCGTGAGGGGCATCATCTTCAGGTGGGAACACCGAG 413

QY 1520 AGAACAAAGG 1528
Db 414 AGAACAAAGG 422

RESULT 6
US-09-864-761-21153
; Sequence 21153, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
```

```
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21153
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121748.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: AW840437.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: O14786, EVALUE 2.00e-23
; OTHER INFORMATION: NT HIT: g11430876, EVALUE 0.00e+00
US-09-864-761-21153

Query Match      8.6%; Score 237.8; DB 10; Length 241;
Best Local Similarity 99.2%; Pred. No. 3.2e-65;
Matches 239; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2521 AGACATCTCAGGAAGCCAGGCAATGTTGAAGACCTTAGAACCCCATCTCATCACC 2580
Db 1 AAGAATCTCCAGGAAGCCAGGCAATGTTGAAGACCTTAGAACCCCATCTCATCACC 60

QY 2581 ATCATAGCCATCAGCCCTGGGGTCTCTCTGGGGGCTGTCTGTGGGCTGTCTGTAC 2640
Db 61 ATCATAGCCATCAGTGCCTGGGGTCTCTCTGGGGGCTGTCTGTGGGCTGTCTGTAC 120

QY 2641 TGTGCTGTGTCATAATGGGATGTGAGAAAGAAACTTCTGTGCCCTGGAGAACTATAAC 2700
Db 121 TGTGCTGTGTCATAATGGGATGTGAGAAAGAAACTTCTGTGCCCTGGAGAACTATAAC 180

QY 2701 TTTGAACCTTGTGGATGGTGTGAAGTTGAAAAAGACAAACTGAATACACAGAGTACTTAT 2760
Db 181 TTTGAACCTTGTGGATGGTGTGAAGTTGAAAAAGACAAACTGAATACACAGAGTACTTAT 240

QY 2761 T 2761
Db 241 T 241

RESULT 7
US-09-867-701-3388
; Sequence 3388, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3388
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(246)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-3388

Query Match      7.8%; Score 215.2; DB 10; Length 246;
Best Local Similarity 96.0%; Pred. No. 4.8e-58;
Matches 217; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 1045 ACCAAGAAGAAATATTATGTCAAGACTTACAAGATCGACGTTAGCTCCAAACGGGAAGAC 1104
Db 7 AGCAAGAAGAAATATTATGTCAAGACTTACAAGATCGACGTTAGCTCCAAACGGGAAGAC 66

QY 1105 TGGATCACCATAAAAAGAGAAACAAACCTGTTCTCTTTTCAGGGAAACACCAACCCACACA 1164
Db 67 TGGTTCCACCATAAAAGAGANGANACAAACCTNTTCTCTTTTCAGGGAAACACCAACCCACACA 126

QY 1165 GATGTTGGTTGCCAGTATTTCCCAACCACTGATAACTCGATTGTCGGAATCAAGCCT 1224
Db 127 GATGTTGGTTGCCAGTATTTCCCAACCACTGATAACTCGATTGTCGGAATCAAGCCT 186

QY 1225 GCAACTTGGGAACTGGCATATCTATGAGATTGGAAGTATACGGTT 1270
Db 187 NCAACTTGGGAACTGGCATATCTATGAGATTGGAAGTATACGGTT 232

RESULT 8
US-09-864-761-18267
; Sequence 18267, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18267
; LENGTH: 195
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121748.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.7
; OTHER INFORMATION: EST_HUMAN HIT: AA976114.1, EVALUE 1.00e-106
; OTHER INFORMATION: SWISSPROT HIT: O14786, EVALUE 2.00e-34
; OTHER INFORMATION: NT HIT: g111430876, EVALUE 1.00e-106
US-09-864-761-18267

Query Match          7.0%; Score 195; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 1e-51;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1420 CCACCGCACCCTCATTCCTACATCAATGAGTGGCTCCAATAGACCTGGGGAGGAGAG 1479
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 CCACCGCACCCTCATTCCTACATCAATGAGTGGCTCCAATAGACCTGGGGAGGAGAG 60

QY 1480 ATCGTGAGGGGCGATCATTCATTCAGGTGGGAAGCACCAGAGAACAAAGGTGTTTCATGAGG 1539
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 ATCGTGAGGGGCGATCATTCATTCAGGTGGGAAGCACCAGAGAACAAAGGTGTTTCATGAGG 120

QY 1540 AAGTTCAAGATCGGTGTACAGCACACACAGCGCTCGGACTGGAAGATGATCATGGATGACAGC 1599
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 AAGTTCAAGATCGGTGTACAGCACACACAGCGCTCGGACTGGAAGATGATCATGGATGACAGC 180

QY 1600 AAACGCAAGCGGAG 1614
      ||||||||||||||
Db 181 AAACGCAAGCGGAG 195

RESULT 9
US-09-864-761-20524
; Sequence 20524, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aomicra-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20524
; LENGTH: 190
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121748.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: NT HIT: g111430876, EVALUE 1.00e-103
; OTHER INFORMATION: EST_HUMAN HIT: R61632.1, EVALUE 8.00e-80
; OTHER INFORMATION: SWISSPROT HIT: O14786, EVALUE 6.00e-34
US-09-864-761-20524

Query Match          6.8%; Score 188.4; DB 10; Length 190;
Best Local Similarity 99.5%; Pred. No. 1.3e-49;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 469 GTGATAAAGTCCCGGATTCCTGAAATATCCCAACAGCCTTGAATGCACCTTATATT 528
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 GTGATAAAGTCCCGGATTCCTGAAATATCCCAACAGCCTTGAATGCACCTTATATT 60

QY 529 GTCTTTGCGCCAAAGATGTCAGAGATTATCCTGGAATTTGAAAGCTTTGACCTGGAGCCT 588
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GTCTTTGCGCCAAAGATGTCAGAGATTATCCTGGAATTTGAAAGCTTTGACCTGGAGCCT 120

QY 589 GACTCAAAATCCTCCAGGGGGGATGTTCTGTCTGCTACGACGGCTAGAAAATCTCGGATGGA 648
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GACTCAAAATCCTCCAGGGGGGATGTTCTGTCTGCTACGACGGCTAGAAAATCTCGGATGGA 180

QY 649 TTCCCTGATG 658
      ||||||||
Db 181 TTCCCTGATG 190

RESULT 10
US-09-864-761-4405
; Sequence 4405, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aomicra-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
```


RESULT 11
US-09-864-761-3757

Query Match	6.1%;	Score 169.2;	DB 10;	Length 389;
Best Local Similarity	93.2%;	Pred. No. 2.8e+43;		
Matches 177;	Conservative	0;	Mismatches 13;	Indels 0;
				Gaps 0;

QY 409 CGTTATGAATTTTCAAGAGAGGTCCTGAATGTTCCAGAACTACACAACACCTACTGGA 468
- - - - -
Db 200 CCGTTTTCTCTCTACTCATAGTCTCTGAATGTTCCAGAACTACACAACACCTACTGGA 259


```

> CURRENT FILING DATE: 2001-05-23
> PRIOR APPLICATION NUMBER: US 60/180,312
> PRIOR FILING DATE: 2000-02-04
> PRIOR APPLICATION NUMBER: US 60/207,456
> PRIOR FILING DATE: 2000-05-26
> PRIOR APPLICATION NUMBER: US 09/632,366
> PRIOR FILING DATE: 2000-08-03
> PRIOR APPLICATION NUMBER: GB 24263.6
> PRIOR FILING DATE: 2000-10-04
> PRIOR APPLICATION NUMBER: US 60/236,359
> PRIOR FILING DATE: 2000-09-27
> PRIOR APPLICATION NUMBER: PCT/US01/006666
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/006667
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/006668
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/006669
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/006670
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/006671
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/006707
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: US 60/234,687
> PRIOR FILING DATE: 2000-09-21
> PRIOR APPLICATION NUMBER: US 09/608,408
> PRIOR FILING DATE: 2000-06-30
> PRIOR APPLICATION NUMBER: US 09/774,203
> PRIOR FILING DATE: 2001-01-29
> NUMBER OF SEQ ID NOS: 49117
> SOFTWARE: Annonax Sequence Listing Engine
> SEQ ID NO 16948
> LENGTH: 241
> TYPE: DNA
> ORGANISM: Homo sapiens
> FEATURE:
> OTHER INFORMATION: MAP TO AC007561.2
> OTHER INFORMATION: EXPRESSED IN HEART
> OTHER INFORMATION: EXPRESSED IN BONE MAR
> OTHER INFORMATION: EXPRESSED IN HBL100,
> OTHER INFORMATION: EXPRESSED IN FETAL L
> OTHER INFORMATION: EXPRESSED IN LUNG. S
> OTHER INFORMATION: EXPRESSED IN ADULT L
> OTHER INFORMATION: EXPRESSED IN PLACENT
> OTHER INFORMATION: EST HIT: g111430587,
> OTHER INFORMATION: EST_HUMAN HIT: AA6833
> OTHER INFORMATION: SWISSPROT HIT: O6046
> US-09-864-761-16948

```

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2002, 20:47:11 ; Search time 2534 Seconds
(without alignments)
17716.610 Million cell updates/sec

Title: US-09-296-264-33

Perfect score: 2772

Sequence: 1 atggagaggggcgctgcgcgt.....gtactatttcggagggcatga 2772

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*

1: em_estba:**
2: em_esthum:**
3: em_estin:**
4: em_estmu:**
5: em_estov:**
6: em_estpl:**
7: em_estro:**
8: em_htc:**
9: gb_est1:**
10: gb_est2:**
11: gb_htc:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estom:**
17: gb_gss:**
18: em_gss_hum:**
19: em_gss_inv:**
20: em_gss_pln:**
21: em_gss_vrt:**
22: em_gss_fun:**
23: em_gss_man:**
24: em_gss_mus:**
25: em_gss_othr:**
26: em_gss_pro:**
27: em_gss_rod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	925	33.4	1024	9 AL573381	AL573381 AL573381
C 2	890.2	32.1	926	9 AL572244	AL572244 AL572244
C 3	848.6	30.6	978	9 AL549032	AL549032 AL549032
C 4	824.8	29.8	988	9 AL547205	AL547205 AL547205
C 5	713.4	25.7	764	14 BM999961	BM999961 UI-H-DP0-
C 6	690	24.9	981	12 BE877043	BE877043 601484847

ALIGNMENTS

RESULT 1
AL573381/C
LOCUS AL573381 1024 bp mRNA linear EST 16-FEB-2001
DEFINITION AL573381 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI042Y013 3
prime, mRNA sequence.
ACCESSION AL573381
VERSION AL573381.1 GI:12932567
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1024)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..1024
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI042Y013"
/clone_lbb="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end

7	682.8	24.6	717	14	BM792307	BM792307 K-EST0072
8	638.4	23.0	640	14	BM792058	BM792058 K-EST0072
9	629.4	22.6	629	9	AL598331	AL598331 DKF2P313J
10	615.6	22.2	803	13	BI822292	BI822292 603036429
11	586.4	21.2	905	14	BQ878747	BQ878747 AGENCOURT
12	579.8	20.9	655	10	BE154199	BE154199 PMI-HT034
13	573.6	20.7	744	14	BQ442311	BQ442311 UI-M-EXO-
14	571.1	20.6	793	13	BI697844	BI697844 603343961
15	570.4	20.6	627	12	BF096148	BF096148 MR3-UT009
16	564	20.3	608	14	BQ016266	BQ016266 UI-H-DT1-
17	557.2	20.1	645	10	AW959505	AW959505 EST371575
18	549	19.8	634	13	BG939321	BG939321 CN31h07.x
19	537.4	19.4	574	9	AA135718	AA135718 zll9c03.f
20	531.4	19.2	608	9	AA156938	AA156938 zll9c03.s
21	516.2	18.6	849	12	BF165167	BF165167 601778032
22	512	18.5	520	10	BE671872	BE671872 7a48f11.x
23	499.8	18.0	673	12	BF466816	BF466816 UI-M-CG0P
24	494.4	17.8	548	10	BE154201	BE154201 PMI-HT034
25	493	17.8	566	10	AV615373	AV615373 AV615373
26	491	17.7	1038	13	BG963500	BG963500 602831167
27	482.6	17.4	617	13	BI738165	BI738165 603362379
28	481.6	17.4	713	12	BF305078	BF305078 601888404
29	476.8	17.2	600	12	BG807721	BG807721 2071-25 M
30	470.4	17.0	869	9	AU035577	AU035577 AU035577
31	470	17.0	535	10	AW103448	AW103448 XG36D05.x
32	467	16.8	604	10	AW489231	AW489231 UI-M-BH3-
33	462.2	16.7	621	10	AW239214	AW239214 XB37d10.y
34	460.8	16.6	476	9	AI652071	AI652071 wb27f03.x
35	457.8	16.5	847	12	BG168336	BG168336 602342307
36	449.6	16.2	916	12	BG327006	BG327006 602428211
37	447.2	16.1	759	9	AU136472	AU136472 AU136472
38	444.8	16.0	779	12	BG120342	BG120342 602353738
39	443.2	16.0	547	9	AA098867	AA098867 zk84f04.s
40	438.8	15.8	471	9	AI139536	AI139536 qc52f11.x
41	436.2	15.7	579	9	AA268329	AA268329 va39b10.f
42	432.4	15.6	543	9	AA099262	AA099262 zk84f04.f
43	431.6	15.6	561	9	AA171321	AA171321 ms50f06.f
44	421.6	15.2	558	14	BM878130	BM878130 IF45D09.y
45	418.6	15.1	545	12	BG086185	BG086185 H3122G06-

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com

BASE COUNT	222 a	257 c	243 g	284 t	18 others
ORIGIN					
Query Match	33.4%; Score 925; DB 9; Length 1024;				
Best Local Similarity	97.7%; Pred. No. 5.2e-246;				
Matches	945; Conservative	11; Mismatches	9; Indels	2; Gaps	2;
QY	882	TTCCCAAGTATGACCAACACTGGTCTGCAGAGCGCTCCCGCTCAACTACCTCAGATGG	941		
Db	1024	TTCCCAAGTATGACCAACACTGGTCTGCAGAGCGCTCCCGCTCAACTACCTCAGATGG	965		
QY	942	GTGAGCTCCCGGAG-AGGATTCCTACCGAGTGGATACAGGTAGACTTGGGCTTCTGC	1000		
Db	964	GTGAGCTCCCGGAGTTCCTACCGAGTGGATACAGGTARACTTGGGCTTCTGC	905		
QY	1001	GCTTTGTACGGCTGTGCGGACACAGGGCGCCATTTCAAAAGAAACCAAGAAATATT	1060		
Db	904	GCTTTGTACGGCTGTGCGGACACAGGGCGCCATTTCAAAAGAAACCAAGAAATATT	845		
QY	1061	ATGCTCAAGACTTACAAGATGACGTTAGTCTCCAAACGGGAAGACTGATCACCATAAAG	1120		
Db	844	ATGCTCAAACTTACAAGATGACGTTAGTCTCCAAACGGGAAGACTGATCACCATAAAG	785		
QY	1121	AAGGAAACAAACCTGTCTCTTTTCAGGGAACACCAACCCACAGATGTTGTGTTGCAG	1180		
Db	784	AAGGAAACAAACCTGTCTCTTTTCAGGGAACACCAACCCACAGATGTTGTGTTGCAG	725		
QY	1181	TATTCGCCAAACACTGATACTCGATTGTCGGAATCAAGCCTGCACTTGGGAACTG	1240		
Db	724	TATTCGCCAAACACTGATACTCGATTGTCGGAATCAAGCCTGCACTTGGGAACTG	665		
QY	1241	GCATATCTATGAGATTGAAGTATACGGTTGCAAGATACAGATTATCTGCTCTGGAA	1300		
Db	664	GCATATCTATGAGATTGAAGTATACGGTTGCAAGATACAGATTATCTGCTCTGGAA	605		
QY	1301	TGTTGGGTATGTHGCTGGAATTTATCTGACCTCCAGATCACATCATCAACCAAGAG	1360		
Db	604	TGTTGGGTATGTHGCTGGAATTTATCTGACCTCCAGATCACATCATCAACCAAGAG	545		
QY	1361	ACAGAACTGGATGCTGGAACATCCGCTGTTAACCAGTCTGCTGCTGGGCACTTC	1420		
Db	544	ACAGAACTGGATGCTGGAATACATCCGCTGTTAACCAGTCTGCTGCTGGGCACTTC	485		
QY	1421	CACCCGCACTCATCTTCACTCAATGAGTGGCTCCAAATAGACCTGGGGAGGAAGA	1480		
Db	484	CACCCGCACTCATCTTCACTCAATGAGTGGCTCCAAATAGACCTGGGGAGGAAGA	425		
QY	1481	TCGTGAGGGGCATCATCTTACGGGTGGGAAGCAGCGAGAACAGAGTGTTCATGAGGA	1540		
Db	424	TCGTGAGGGGCATCATCTTACGGGTGGGAAGCAGCGAGAACAGAGTGTTCATGAGGA	365		
QY	1541	AGTTCAAGATCGGGTACACACACAGGGCTCGGACTGGAAGATGATCATGATGACAGCA	1600		
Db	364	AGTTCAAGATCGGGTACACACACAGGGCTCGGACTGGAAGATGATCATGATGACAGCA	305		
QY	1601	AACGCAAGCGGAAGTCTTTTGGGGCAACAACATATCATACCTGAGCTGCGGACTT	1660		
Db	304	AACGCAAGCGGAAGTCTTTTGGGGCAACAACATATCATACCTGAGCTGCGGACTT	245		
QY	1661	TTCCAGCTCTCCACGGGATTTCATCAGGATCTACCCGAGAGAGCCATCATGCGGAC	1720		
Db	244	TTCCAGCTCTCTCCACGGGATTTCATCAGGATCTACCCGAGAGAGCCATCATGCGGAC	185		
QY	1721	TCGGGCTCAGATGAGCTGCTGGGCTGTGAAGTGGAAACCCCTACAGCTG-GACCGACC	1779		

Db	184	TGGGGCTAGAAATGAGCTGCTGGGCTGTGAAGTGGAAAGCCCTTACARMTGMGACCGACR	125		
Qy	1780	ACTCCCAACGGGAACCTTGGTGGATGAATGTGATGACACAGGCAACTGCCACAGTGA	1839		
		I:::			
Db	124	AVTSCCAACGGGAACCTTGGTGGATGAATGTGATGACACAGGCAACTGCCACAGTGGN	65		
Qy	1840	ACAGGTG	1846		
Db	64	ACAGGTG	58		
RESULT 2					
AL572244/c					
LOCUS					
DEFINITION	AL572244	LTI_NFL006_PL2	926 bp	mRNA	linear
	AL572244	LTI_NFL006_PL2	Homo sapiens	cDNA clone	CSODI010YF07 3
ACCESSION	AL572244	prime, mRNA sequence.			
VERSION	AL572244.1	GI:12930323			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.				
FEATURES					
source	1..926				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="CSODI010YF07"				
	/clone_lib="LTI_NFL006_PL2"				
	/tissue_type="placenta"				
	/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"				
BASE COUNT	198 a	237 c	225 g	261 t	5 others
ORIGIN					
Query Match	32.1%; Score 890.2; DB 9; Length 926;				
Best Local Similarity	97.5%; Pred. No. 2.4e-236;				
Matches	898; Conservative	5; Mismatches	18; Indels	0; Gaps	0;
QY	858	TCATTCTGACAGATCACAGCTTCTTCCAGTATAGCACTGGTCTGCAGAGCGCTC	917		
Db	926	TCATTCTGACAGATCACAGCTTCTTCCAGTATAGCACTGGTCTGCAGAGCGCTC	867		
QY	918	CCGCCTGAACCTACCGTGGAGATGGGTGGACTCCCGGAGAGGATTCCTACCGAGAGTGGAT	977		
Db	866	CCGCCTGAACCTACCGTGGAGATGGGTGGACTCCCGGAGAGGATTCCTACCGAGAGTGGAT	807		
QY	978	ACAGGTAGACTTGGGCTTCTGCGCTTGTTCAGGGCTGTGCGGACACAGGGCGCCATTTC	1037		
Db	806	ACAGGTAGACTTGGGCTTCTGCGCTTGTTCAGGGCTGTGCGGACACAGGGCGCCATTTC	747		
QY	1038	AAAAGAAACCAAGAAATATTTATGTCAGACTTACAAGATCGAGTTAGCTTCAACGG	1097		
Db	746	AAAAGAAACCAAGAAATATTTATGTCAGACTTACAAGATCGAGTTAGCTTCAACGG	687		
QY	1098	GGAAGACTGGATCACCATAAAGAGAAACAACTGTTCTTTTCAGGGAAACACCAA	1157		

Db	686	GAAGACTGGATCACCATATAAAGAAAGNAACAAACACTGTTCTCTTCAGGGAAACACCAA	627
Qy	1158	CCCACAGAGATGTTGTGGTTGCAGTATTCCCCAAACCACTGATAACTCGATTGTGCGGAAT	1217
Db	626	CCCTACAGATGAAAWAATTTGCAGTATTTCCCAACCACTGATAACTCGATTGTGCGGAAT	567
Qy	1218	CAAGCCTGCAACTTTGGGAAACTGGCATATCTATGAGATTTGAAGTATATGTTTGAAGAT	1277
Db	566	CAAGCCTGCAACTTTGGGAAACTGGCATATCTATGAGATTTGAAGTATATGTTTGAAGAT	507
Qy	1278	AACAGATTATCTTGTCTCTGGAATGTTGGGTATGGTGTCTGAGCTTATTTCTGACTCCCA	1337
Db	506	AACAGATTATCTTGTCTCTGGAATGTTGGGTATGGTGTCTGAGCTTATTTCTGACTCCCA	447
Qy	1338	GATCACATCATCCAAACCAAGGAGACAGAAACTGGATGGCTGAAAAACATCCGCCCTGGTAAC	1397
Db	446	GATCACATCATCCAAACCAAGGAGACAGAAACTGGATGGCTGAAAAACATCCGCCCTGGTAAC	387
Qy	1398	CAGTCGCTCTGGCTGGGCACTTCCACCGGACCTCATCTTCCATACATCAATGAGTGGCTCCA	1457
Db	386	CAGTCGCTCTGGCTGGGCACTTCCACCGGACCTCATCTTCCATACATCAATGAGTGGCTCCA	327
Qy	1458	AATAGACTCTGGGAGGAGAGATGCTGAGGGGATCATCATTCAGGCTGGGAGGAGACCG	1517
Db	326	AATAGACTCTGGGAGGAGAGATGCTGAGGGGATCATCATTCAGGCTGGGAGGAGACCG	267
Qy	1518	AGAGAACAAAGTGTTCATGAGAAAGTTCAAGATCGGGTACAGCAACACGGCTCGGACTG	1577
Db	266	AGAGAACAAAGTGTTCATGAGAAAGTTCAAGATCGGGTACAGCAACACGGCTCGGACTG	207
Qy	1578	GAAGATGATCATGATGATGACAGCAACGCAAGCGGAAAGTCTTTTGGAGGCAACACAACTA	1637
Db	206	GAAGATGATCATGATGATGACAGCAACGCAAGCGGAAAGTCTTTTGGAGGCAACACAACTA	147
Qy	1638	TGATACACTCAGCTCGGCACTTTTCCAGCTCTCTCCACGCGATTCATCAGGATCTACCC	1697
Db	146	TGATACACTCAGCTCGGCACTTTTCCAGCTCTCTCCACGCGATTCATCAGGATCTACCC	87
Qy	1698	CGAGAGACCCACTCATGGCGACTGGGGCTCAGAAATGGAGCTGCTGAGGTGGA	1757
Db	86	CGAGAGACCCACTCATGGCGACTGGGGCTCAGAAATGGAGCTGCTGAGGTGGA	27
Qy	1758	AGCCCTTACACTGGACCGAC	1778
Db	26	AGGTGGCACCANTGTGTGC	6
RESULT 3			
AL549032			
LOCUS	AL549032	LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI042Y013	5
DEFINITION	AL549032	prime, mRNA sequence.	
ACCESSION	AL549032		
VERSION	AL549032.1	GI:12884623	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 978)		
JOURNAL	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.		
COMMENT	Full-length cDNA libraries and normalization		
	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
FEATURES	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
source	Location/Qualifiers		
	1..978		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DI042Y013"		
	/clone_lib="LTI_NFL006_PL2"		

/note="Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI_CGAP_Fsl is a cDNA library containing the following tissue(s): Fibrosarcoma Cell line HT-1080 (ATCC number CCL-121). The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is GTTCTACGAG.

TAG_LIB=UI-H-DPO
TAG_TISSUE=fibrosarcoma
TAG_SEQ=GTCTACGAG"

BASE COUNT 150 a 200 c 183 g 229 t 2 others
ORIGIN

Query Match 25.7%; Score 713.4; DB 14; Length 764;
Best Local Similarity 99.6%; Pred. No. 4e-187;
Matches 714; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1208 TTGTCGGAATCAAGCTCGAATCTGGAACTGGGATATCTATGAGATTTGAAGTATAGC 1267
Db 761 TTGTCGGAATCAAGCTCGAATCTGGAACTGGGATATCTATGAGATTTGAAGTATAGC 702
QY 1268 GTTGAAGATACAGATATCTTCTGCTGGAATGTTGGGTATGTTGCTGACATATTT 1327
Db 701 GTTGAAGATACAGATATCTTCTGCTGGAATGTTGGGTATGTTGCTGACATATTT 642
QY 1328 CTGACTCCAGATACATCATCAACACGAGGAGACAGAACTGGATGCTGAAACATCC 1387
Db 641 CTGACTCCAGATACATCATCAACACGAGGAGACAGAACTGGATGCTGAAACATCC 582
QY 1388 GCCTGGTAACCAAGTCTGCTGGGCACTTCCACCGCACCTCATCTCATCAATG 1447
Db 581 GCCTGGTAACCAAGTCTGCTGGGCACTTCCACCGCACCTCATCTCATCAATG 522
QY 1448 ACTGGCTCCAAATAGACCTGGGGAGGAGAGATCGTAGGGGCGATCATCTCAGGTG 1507
Db 521 ACTGGCTCCAAATAGACCTGGGGAGGAGAGATCGTAGGGGCGATCATCTCAGGTG 462
QY 1508 GGAAGCACCAGAGAACAAAGTGTTCATGAGGAAGTTCAAGATCGGTACACCAACAG 1567
Db 461 GGAAGCACCAGAGAACAAAGTGTTCATGAGGAAGTTCAAGATCGGTACACCAACAG 402
QY 1568 GCTCGAGCTGGAAGATGATCATGATGATGATGATGATGATGATGATGATGATGAT 1627
Db 401 GCTCGAGCTGGAAGATGATCATGATGATGATGATGATGATGATGATGATGATGAT 342
QY 1628 ACAACAACTATCATACACTGAGCTGGGACCTTTTCCAGCTCTCTCCAGCGCATCATCA 1687
Db 341 ACAACAACTATGATACACTGAGCTGGGACCTTTTCCAGCTCTCTCCAGCGCATCATCA 282
QY 1688 GGATCTACCCCGAGAGAGCCACTCATGCGGACTGGGGCTCAGAGTGGAGCTGCTGGGCT 1747
Db 281 GGATCTACCCCGAGAGAGCCACTCATGCGGACTGGGGCTCAGAGTGGAGCTGCTGGGCT 222
QY 1748 GTGAAGTGAAGCCCTTACAGCTGGACCGACCACTCCCAACGGGAACTTGTGGATGAAT 1807
Db 221 GTGAAGTGAAGCCCTTACAGCTGGACCGACCACTCCCAACGGGAACTTGTGGATGAAT 162
QY 1808 GTGATGAGCAGCAGCCCACTCCACAGTGGAAAGGTGATGACTTCCAGCTTCAGGTG 1867
Db 161 GTGATGAGCAGCAGCCCACTCCACAGTGGAAAGGTGATGACTTCCAGCTTCAGGTG 102
QY 1868 GCACCACTGTGTGTCACAGAAAGCCACCGTTCATAGACAGCACCACATACAAATCAG 1924
Db 101 GCACCACTGTGTGTCACAGAAAGCCACCGTTCATAGACAGCACCACATACAAATCAG 45

RESULT 6
LOCUS BE877043
DEFINITION 601484847F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3887435 5', mRNA sequence.
ACCESSION BE877043
VERSION BE877043.1 GI:10325819
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 981)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9665 row: n column: 12
High quality sequence stop: 646.
FEATURES
Location/Qualifiers
1..981
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3887435"
/clone_lib="NIH_MGC_69"
/tissue_type="large cell carcinoma, undifferentiated"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."
BASE COUNT 287 a 228 c 270 g 196 t
ORIGIN
Query Match 24.9%; Score 690; DB 12; Length 981;
Best Local Similarity 96.0%; Pred. No. 1.5e-180;
Matches 751; Conservative 0; Mismatches 25; Indels 6; Gaps 4;
QY 1861 ACAGTGGCACACATGCTGCTGCCACAGAAAAGCCACGGTCATAGACAGCACCATACAA 1920
Db 1 ACAGTGGCACACATGCTGCTGCCACAGAAAAGCCACGGTCATAGACAGCACCATACAA 60
QY 1921 TCAGAGTTTCCAAACATATGTTTAACTGTGAATTTGGCTGGGCTCTCCAAAGACCTTC 1980
Db 61 TCAGAGTTTCCAAACATATGTTTAACTGTGAATTTGGCTGGGCTCTCCAAAGACCTTC 120
QY 1981 TGCCACTGGGAACATGACAATCACCTGCAGCTCAAGTGGAGTGTGTTCACCAAGACAGC 2040
Db 121 TGCCACTGGGAACATGACAATCACCTGCAGCTCAAGTGGAGTGTGTTCACCAAGACAGC 180
QY 2041 GGACCCATTCAGGATCACACAGAGATGGCAACTTCATCTATTCCCAAGCTCACGAAAT 2100
Db 181 GGACCCATTCAGGATCACACAGAGATGGCAACTTCATCTATTCCCAAGCTCACGAAAT 240
QY 2101 CAGAAAGGCAAGTGGCTGCGCTGGTGGAGCCCTGTTTATTTCCAGAACTCTGCCAC 2160
Db 241 CAGAAAGGCAAGTGGCTGCGCTGGTGGAGCCCTGTTTATTTCCAGAACTCTGCCAC 300
QY 2161 TGCATGACCTTCTGGTATCATCATGTCTGGGTCCCGCTGGCAGCTCAGGGTCAAACTG 2220
Db 301 TGCATGACCTTCTGGTATCATCATGTCTGGGTCCCGCTGGCAGCTCAGGGTCAAACTG 360
QY 2221 CGCTACCAAGACCCAGAGAGTACGATCAGCTGGTGTCTGGATGCCATTGGACACCAAGGT 2280
Db 361 CGCTACCAAGACCCAGAGAGTACGATCAGCTGGTGTCTGGATGCCATTGGACACCAAGGT 420

```
QY 2281 GACCACCTGGAAGGAGGCGTCTTCTGCTCCACAAGTCTCTGAAACTTTATCAGGTGATT 2340
Db 421 GACCACCTGGAAGGAGGCGTCTTCTGCTCCACAAGTCTCTGAAACTTTATCAGGTGATT 480
QY 2341 TTCCAGGGCCAAATCGGAAAGAAACCTTGGTGGGATGCTGTGGATGACATAGTATT 2400
Db 481 TTCCAGGGCCAAATCGGAAAGAAACCTTGGTGGGATGCTGTGGATGACATAGTATT 540
QY 2401 AATAAACACATTTTCAACAAGAGATTGTGCAAAACACGACAGACCTGGATAAAAAGAACCCA 2460
Db 541 AATAAACACATTTTCAACAAGAGATTGTGCAAAACACGACAGACCTGGATAAAAAGAACCCA 600
QY 2461 GAAATTAATAATGATGAACAGGAGACACGCCAGGATACGAAGGTGAAGAGAGAGGTGAC 2520
Db 601 GAAATT-AAATTGATGAACAGGAGACACGCCAAGATACGAAGGTGAAGAGAGCAAGGTGA 659
QY 2521 AAGAACATCTCC--AGGAAGCCAGGCAAAAGTGTGTGA--GACCTTAGAACCCCATCTCAT 2576
Db 660 CAGAAGCTCTCCAGGAAAGCCAGGCAATGTGTGAAGACCTTTAGAACCCCATCTCAT 719
QY 2577 CACCATCATAGCCATGAGCCCTCGGGGTCTCTCT--GGGGGCTGTCTGTGGGGTCTGTGC 2635
Db 720 CACATCATAGCCATTAGTGCCTCGGGGTCTCTCTGGGGGGTGGCTGTGGGGTCTGGGC 779
QY 2636 TG 2637
Db 780 TG 781

RESULT 7
BM792307
LOCUS
DEFINITION K-EST0072500 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-13-A12
5', mRNA sequence.
ACCESSION BM792307
VERSION BM792307.1 GI:19140539
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 717)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 13 row: A column: 12
High quality sequence stop: 717.
FEATURES
Location/Qualifiers
1..717
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S22SNU16n1-13-A12"
/clone_lib="S22SNU16n1"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/lab_host="SNU-16"
/note="Organ: Stomach; Vector: pT7T3-Pac; Site:1: EcoRI;
Site:2: NotI; The S22SNU16 library was contributed by the
Soares laboratory and it was constructed as described by
Ronald, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. RNA was prepared from harvested
```

```
cells of SNU-16 culture. SNU-16 cell was obtained from
Korean Cell Line Bank (KCLB). SNU-16 was established from
ascitic fluids of Korean patients by Park J.G. et al.
(1990), Cancer Res 50: 2773-2780."
BASE COUNT 200 a 182 c 190 g 145 t
ORIGIN
Query Match 24.6%; Score 682.8; DB 14; Length 717;
Best Local Similarity 96.9%; Pred. No. 1.3e-178;
Matches 715; Conservative 0; Mismatches 2; Indels 21; Gaps 1;
QY 1185 CCCAAACCACTGATTAACCTGATTTGTCGAATCAAGCCTGCAACTTTGGAAACTTGCAT 1244
Db 1 CCCAAACCACTGATTAACCTGATTTGTCGAATCAAGCCTGCAACTTTGGAAACTTGCAT 60
QY 1245 ATCTATGAGATTGAAGTATACGTTTGCAGATAACAGATTATCTTGCCTCTGGAATGTT 1304
Db 61 ATCTATGAGATTGAAGTATATGTTGCAAGATAACAGATTATCTTGCCTCTGGAATGTT 120
QY 1305 GGGTATGGTCTCTGGACTTATTTCTGACTCCAGATCACATCATCAACCAAGAGACAG 1364
Db 121 GGGTATGGTCTCTGGACTTATTTCTGACTCCAGATCACATCATCAACCAAGAGACAG 180
QY 1365 AACTGTGATGCCCTGAAAACATCCGCTGTTAAACAGTCGCTCTGGCTGGCATTCCACC 1424
Db 181 AACTGTGATGCCCTGAAAACATCCGCTGTTAAACAGTCGCTCTGGCTGGCATTCCACC 240
QY 1425 CGCACCTCATCTTACATCAATCAGTGGCTCCAAATAGACTCTGGGGAGGAGAGATCGT 1484
Db 241 CGCACCTCATCTTACATCAATCAGTGGCTCCAAATAGACTCTGGGGAGGAGAGATCGT 300
QY 1485 GAGGGCATCATCATTTCAGGGTGGGAAGCACCAGAGAGAACAAGGTGTTTCATGAGGAAGTT 1544
Db 301 GAGGGCATCATCATTTCAGGGTGGGAAGCACCAGAGAGAACAAGGTGTTTCATGAGGAAGTT 360
QY 1545 CAAGATCGGGTACAGCAACACCGCTCGGACTCGAGATGATCATGATGACAGCAACAG 1604
Db 361 CAAGATCGGGTACAGCAACACCGCTCGGACTCGGAGATGATCATGATGACAGCAACAG 420
QY 1605 CAAGGCGAAGTCTTTGAGGGCAACAACACTATGATACACCTGAGCTGGGACTTTTCC 1664
Db 421 CAAGGCGAAGTCTTTGAGGGCAACAACACTATGATACACCTGAGCTGGGACTTTTCC 480
QY 1665 AGCTCTCTCCACGCGATTTCATCAGGATCTACCCCGAGAGAGCCACTCATGGCGGACTGGG 1724
Db 481 AGCTCTCTCCACGCGATTTCATCAGGATCTACCCCGAGAGAGCCACTCATGGCGGACTGGG 540
QY 1725 GCTCAGATGGAGCTGCTGGGCTGTGAAGTGGAGCCCTACAGCTGGACCGACCTCC 1784
Db 541 GCTCAGATGGAGCTGCTGGGCTGTGAAGTGGAGCCCTACAGCTGGACCGACCTCC 600
QY 1785 CAACGGGAACTTGGTGGATGAATGTGATGACGACGAGGCAACTGCCACACTGGAACAGG 1844
Db 601 CAACGGGAACTTGGTGGATGAATGTGATGACGACGAGGCAACTGCCACACTGGAACAGG 655
QY 1845 TGATGACTTCCAGCTCAGAGTGGCACCACCTGTGCTGGCCACAGAAAAGCCACGGTCAT 1904
Db 656 -----ACAGGTGGCACCCTGTGCTGGCCACAGAAAAGCCACGGTCAT 699
QY 1905 AGACAGCACCATACAATC 1922
Db 700 AGACAGCACCATACAATC 717

RESULT 8
BM792058
LOCUS
DEFINITION K-EST0072198 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-10-H10
5', mRNA sequence.
ACCESSION BM792058
VERSION BM792058.1 GI:19140290
KEYWORDS EST.
SOURCE human.
```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 10 row: H column: 10
High quality sequence scop: 640.
FEATURES
source
Location/Qualifiers
1..640
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S22SNU16n1-10-H10"
/clone_lib="S22SNU16n1"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/lab_host="DH10B"
/note="Organ: Stomach; Vector: pT73-Pac; Site_1: EcoRI;
Site_2: NotI; The S22SNU16 library was contributed by the
Soares laboratory and it was constructed as described by
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. RNA was prepared from harvested
cells of SNU-16 culture. SNU-16 cell was obtained from
Korean Cell Line Bank (KCLB). SNU-16 was established from
ascitic fluids of Korean patients by Park J.G. et al.
(1990), Cancer Res 50: 2773-2780."
BASE COUNT 188 a 154 c 155 g 143 t
ORIGIN
Query Match 23.0%; Score 638.4; DB 14; Length 640;
Best Local Similarity 99.8%; Pred. No. 2.9e-166;
Matches 639; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1058 ATTATGTCAAGACTTACAGATCGAGTTAGCTCTTTCAGGGAAACACCAACCCGGAAGACTGGATCACCATAA 1117
Db 1 ATATGTCAGACTTTACAGATCGAGTTAGCTCTTTCAGGGAAACACCAACCCGGAAGACTGGATCACCATAA 60

QY 1118 AAGAAGAAACAAACCTGTCTCTTTTCAGGGAAACACCAACCCGGAAGACTGGATGTTGGTTG 1177
Db 61 AAGAAGAAACAAACCTGTCTCTTTTCAGGGAAACACCAACCCGGAAGACTGGATGTTGGTTG 120

QY 1178 CAGTATTCCTCCAAACCACTGATACTGATTCGATTGTTCGGAATCAAGCTGCAACTGGGAAA 1237
Db 121 CAGTATTCCTCCAAACCACTGATACTGATTCGATTGTTCGGAATCAAGCTGCAACTGGGAAA 180

QY 1238 CTGGCATATCTATGAGATTGGAATGATACGGTTGCAAGATACACATATATCCTTCTGCTG 1297
Db 181 CTGGCATATCTATGAGATTGGAATGATACGGTTGCAAGATACACATATATCCTTCTGCTG 240

QY 1298 GAATGTTGGGTATGGTGTCTGGACTATTTCCTGACTCCGAGATCATCATCCAAACCAAG 1357
Db 241 GAATGTTGGGTATGGTGTCTGGACTATTTCCTGACTCCGAGATCATCATCCAAACCAAG 300

QY 1358 GAGACAGAACTGGATGCTTGAACATCCGCTGGTACCAAGTCGCTGCTGGGAC 1417
Db 301 GGGACAGAACTGGATGCTTGAACATCCGCTGGTACCAAGTCGCTGCTGGGAC 360

QY 1418 TTCCACCGCACCCTCATTCCTACATCAATGAGTGTGCTCCAAATAGACCTGGGGAGGAGA 1477
Db 361 TTCCACCGCACCCTCATTCCTACATCAATGAGTGTGCTCCAAATAGACCTGGGGAGGAGA 420

```

```

QY 1478 AGATCGTGAAGGGCATCATTCATTCAGGTGGGAAGCACCAGAGAACAAAGGTGTTTCATGA 1537
Db 421 AGATCGTGAAGGGCATCATTCATTCAGGTGGGAAGCACCAGAGAACAAAGGTGTTTCATGA 480

QY 1538 GGAAGTTCAAGATCGGGTACACACCAACCGCTCGGACTGGGAAGATGATCATGGATGACA 1597
Db 481 GGAAGTTCAAGATCGGGTACACACCAACCGCTCGGACTGGGAAGATGATCATGGATGACA 540

QY 1598 GCAAAACGCAAGCGGAAGTCTTTTGGGGCAACCAACACTATGATACACCTGAGCTGCGGA 1657
Db 541 GCAAAACGCAAGCGGAAGTCTTTTGGGGCAACCAACACTATGATACACCTGAGCTGCGGA 600

QY 1658 CTTTTCAGCTCTCTCCACGCGATTTCATCAGGATCTACCC 1697
Db 601 CTTTTCAGCTCTCTCCACGCGATTTCATCAGGATCTACCC 640

RESULT 9
AL598331 629 bp mRNA linear EST 14-AUG-2001
LOCUS DKFZp313J041_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
DEFINITION DKFZp313J041 5', mRNA sequence.
ACCESSION AL598331
VERSION AL598331.1 GI:15161022
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 629)
Poustka,A., Wellenreuther,R., Meves,H.W., Weil,B. and Wiemann,S.).
EST (Poustka,A., Wellenreuther,R., Meves,H.W., Weil,B. and Wiemann,S.).
JOURNAL Unpublished (1999)
COMMENT Contact: Poustka A.J.
Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestrasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mpg-berlin-dahlem.mpg.de
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp313J041) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
Location/Qualifiers
1..629
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp313J041"
/clone_lib="313 (synonym: hlcc2)"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: pTriplex2; Site_1: sf1a; Site_2: sf11b;
cDNA-collection"
BASE COUNT 168 a 154 c 151 g 156 t
ORIGIN
Query Match 22.6%; Score 627.4; DB 9; Length 629;
Best Local Similarity 99.8%; Pred. No. 3.3e-163;
Matches 628; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 376 GACTACGAAACACATGGTGCAGGATTTTCCATACCTTATGAATTTTCAAGAGAGTCTCT 435
Db 1 GACTACGAAACACATGGTGCAGGATTTTCCATACCTTATGAATTTTCAAGAGAGTCTCT 60

QY 436 GAATGTTCCCAACTACACACACCTAGTGGAGTGATAAAGTCCCGGAGTTCCTCTGAA 495

```

Db 61 GAATGTTCCAGACATACACACCTAGTGGAGTGATAAAGTCCCGGATTCCCTGAA 120
QY 496 AAATATCCCAACAGCCTTGAATCACTATATATGCTTTGCGCCAAAGATGTCCAGAGATT 555
Db 121 AAATATCCCAACAGCCTTGAATCACTATATATGCTTTGCGCCAAAGATGTCCAGAGATT 180
QY 556 ATCCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAAATCCTCCAGGGGGATGTTT 615
Db 181 ATCCTGGAATTTGAAAGCTTTGACCTGGAGCCTGACTCAAAATCCTCCAGGGGGATGTTT 240
QY 616 TGTGCGTACGACCGGCTAGAAATCTGGGATGGAATTCCTGATGTTGGCCCTCACATTGG 675
Db 241 TGTGCGTACGACCGGCTAGAAATCTGGGATGGAATTCCTGATGTTGGCCCTCACATTGG 300
QY 676 CGTTACTGTGGACAGAAAACACCAAGTTCGAAATCCGATCCGATCCTATCCTGCGCATCTCTCCATG 735
Db 301 CGTTACTGTGGACAGAAAACACCAAGTTCGAAATCCGATCCTATCCTGCGCATCTCTCCATG 360
QY 736 GTTTTTTACACGACGCGGATAGCAAAAGAGTTTCTCAGCAAACTACAGTGTCTTG 795
Db 361 GTTTTTTACACGACGCGGATAGCAAAAGAGTTTCTCAGCAAACTACAGTGTCTTG 420
QY 796 CAGAGCAGTCTCAGAGATTTCAATGTATGGAAGCTCTGGGCATGGAATCAGAGAA 855
Db 421 CAGAGCAGTCTCAGAGATTTCAATGTATGGAAGCTCTGGGCATGGAATCAGAGAA 480
QY 856 ATTCATTTTACACGATCAGAGCTTCTCCAGTATAGCAACCACTGGTCTGAGAGCGC 915
Db 481 ATTCATTTTACACGATCAGAGCTTCTCCAGTATAGCAACCACTGGTCTGAGAGCGC 540
QY 916 TCCCGCTGAACCTACCTGAGATGGTGGACTCCCGAGAGGATTCCTACCGAGAGTGG 975
Db 541 TCCCGCTGAACCTACCTGAGATGGTGGACTCCCGAGAGGATTCCTACCGAGAGTGG 600
QY 976 ATCAGGTAGACTGGGCTTCTGCGCTT 1004
Db 601 ATCAGGTAGACTGGGCTTCTGCGCTT 629

RESULT 10
BI822292
LOCUS 603036429F1 NIH_MGC_115 Homo sapiens cDNA clone image:5177551 5',
DEFINITION mRNA sequence.
ACCESSION BI822292
VERSION BI822292.1 GI:15933842
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 803)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L14M1442 row: i column: 08
High quality sequence stop: 801.
Location/Qualifiers
1. .803
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5177551"
/clone_lib="NIH_MGC_115"

FEATURES
source

RESULT 11
BQ878747
LOCUS
DEFINITION
ACCESSION

BQ878747 905 bp mRNA linear EST 16-AUG-2002
AGENCOURT_7980968 Lupski_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6185475 5', mRNA sequence.
BQ878747

/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 193 a 202 c 193 g 215 t
ORIGIN
Query Match 22.2%; Score 615.6; DB 13; Length 803;
Best Local Similarity 98.2%; Pred. No. 7.3e-160;
Matches 654; Conservative 0; Mismatches 9; Indels 3; Gaps 3;
QY 1 ATGAGAGGGGGCTGCGGCTCTCTGCGGCTGCTCGCCCTCGCTCGCCCGCGCGGC 60
Db 139 ATGAGAGGGGGCTGCGGCTCTCTGCGGCTGCTCGCCCTCGCTCGCCCGCGCGGT 198
QY 61 GCTTTTCGCAACGATGAATGTGCGGATCTATAAAATTTGAAAGCCCGGGTACCTTACA 120
Db 199 GCTTTTCGCAACGATGAATGTGCGGATCTATAAAATTTGAAAGCCCGGGTACCTTACA 258
QY 121 TCTCCTGGTTATCTCTCATCTTATCACCCCAAGTGAATAATGCGAATGGCTTCAAGCT 180
Db 259 TCTCCTGGTTATCTCTCATCTTATCACCCCAAGTGAATAATGCGAATGGCTTCAAGCT 318
QY 181 CCGGACCCATACAGAGAAATATGATCACTTCAACCTCAGTTCGATTTGGAGGACAGA 240
Db 319 CCGGACCCATACAGAGAAATATGATCACTTCAACCTCAGTTCGATTTGGAGGACAGA 378
QY 241 GACTGCAAGTATCAGTACGTTGAAGTCTTCGATGGAGAAATGAAATGACATTTTAGG 300
Db 379 GACTGCAAGTATCAGTACGTTGAAGTCTTCGATGGAGAAATGAAATGACATTTTAGG 438
QY 301 GGAAGTTCTGTGGAAGATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTTT 360
Db 439 GGAAGTTCTGTGGAAGATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTTT 498
QY 361 ATCAAAATTTGCTCTGACTACGAAACACATGTTGCGAGGATTTTCCATACATTTTAAAT 420
Db 499 ATCAAAATTTGCTCTGACTACGAAACACATGTTGCGAGGATTTTCCATACATTTTAAAT 558
QY 421 TTCAAGAGAGGCTCTGAAATGTTCCAGAACTACACACACCTAGTGAGTATAAAGTCC 480
Db 559 TTCAAGAGAGGCTCTGAAATGTTCCAGAACTACACACACCTAGTGAGTATAAAGTCC 617
QY 481 CCGGATTTCCCTGAAATAATCCCAACAGCCTTTGAATGCACTTATATTTGTCGCGCA 540
Db 618 CCGGATTTCCCTGAAATAATCCCAACAGCCTTTGAATGCACTTATATGTTGCTTTGCGCCA 677
QY 541 AGATGTCAGAGATTTATCCTGGAATTTGAAGC-TTTGACCTGAGCCTGACTCAATTC 599
Db 678 AGATGTCAGAGATTTATCCTGGAATTTGAAGC-TTTGACCTGAGCCTGACTCAATTC 737
QY 600 TCCAG-GGGGGATGTTCTGTCGTACGCGCTAGACAAATCTGGGATGATTCCTGTATG 658
Db 738 TCCAGGGGGGATGTTCTGTCGTAAACGCGGTAGAAATCTGGGATGATTCCTGTATG 797
QY 659 TTGGCC 664
Db 798 TTGGCC 803

```

VERSION B0878747.1 GI:22270755
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rcpbbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13576 row: f column: 04
High quality sequence stop: 605.
FEATURES
Source
Location/Qualifiers
1..905
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6185475"
/clone_lib="Lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
Not1; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCTACGCGTCCG-3' and
5'-GACTAGTTCTAGTCGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT 255 a 186 c 250 g 210 t 4 others
ORIGIN
Query Match 21.2%; Score 586.4; DB 14; Length 905;
Best Local Similarity 99.8%; Pred No. 1.1e-151;
Matches 587; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2185 TCTGGTCCCGTCCGACACTCAGGTCGCTCAAACTCGGCTACCAAGCCAGAGGAGTAC 2244
Db 1 TCTGGTCCCGTCCGACACTCAGGTCGCTCAAACTCGGCTACCAAGCCAGAGGAGTAC 60
QY 2245 GATCAGCTGGTCTGATGCCATTTGGACACCAAGGTGACCATGGAAGGAGGGGTGTC 2304
Db 61 GATCAGCTGGTCTGATGCCATTTGGACACCAAGGTGACCATGGAAGGAGGGGTGTC 120
QY 2305 TTGCTCCACAACTCTGAACTTTATCAGGTGATTTTCGAGGCGGAATCGGAAAGGA 2364
Db 121 TTGCTCCACAACTCTGAACTTTATCAGGTGATTTTCGAGGCGGAATCGGAAAGGA 180
QY 2365 AACCTTGGTGGATTGCTGTGGATGACATTTAGTATTAATCAACATTTTCAACAAGAGT 2424
Db 181 AACCTTGGTGGATTGCTGTGGATGACATTTAGTATTAATCAACATTTTCAACAAGAGT 240
QY 2425 TGTGCAAAACACGACAGCTGTGATAAAAGAACCCAGAAATTAATTAATGATGAACAGGG 2484
Db 241 TGTGCAAAACACGACAGCTGTGATAAAAGAACCCAGAAATTAATTAATGATGAACAGGG 300
QY 2485 AGCAGCCAGGATACGAGGTCAAGGAGTGAAGGTGACAAACATCTCCAGGAGCCAGGC 2544
Db 301 ACACGCCAGGATACGAGGTGAAGGAGTGAAGGTGACAAACATCTCCAGGAGCCAGGC 360
QY 2545 AATGTGTTGAAGACCTTAGAACCCCTCTCATCACCATCATAGCCATGAGCGCCCTGGGG 2604

```

```

Db 361 AATGTGTTGAAGACCTTAGACCCCATCTCATCACCATCATAGCCATGAGCCCTGGGG 420
QY 2605 GTCCCTCCTGGGGGCTGTCTGTGGGCTGTGCTGTACTGTGCTGTGGCATATATGGGATG 2664
Db 421 GTCCCTCCTGGGGGCTGTCTGTGGGCTGTGCTGTACTGTGCTGTGGCATATATGGGATG 480
QY 2665 TCAGAAAGAAACTTCTGCGCCCTGGAGAACTATAACTTTGAACTTGTGATGCTGTGAAG 2724
Db 481 TCAGAAAGAAACTTCTGCGCCCTGGAGAACTATAACTTTGAACTTGTGATGCTGTGAAG 540
QY 2725 TTGAAAAAGACAAACTGAATACACAGAGTACTTATTCGGAGGCATGA 2772
Db 541 TTGAAAAAGACAAACTGAATACACAGAGTACTTATTCGGAGGCATGA 588
RESULT 12
BE154199
LOCUS PM1-HT0340-150300-005-b08 HT0340 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE154199
ACCESSION BE154199
VERSION BE154199.1 GI:8616920
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 655)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Negai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-PM1-HT0340-150
300-005-b08&t3=2000-03-15&t4=1)
Seg primer: puc 18 forward
High quality sequence stop: 618.
FEATURES
Location/Qualifiers
source
1..655
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0340"
/dev_stage="Adult"
/notes="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 182 a 148 c 181 g 143 t 1 others
ORIGIN
Query Match 20.9%; Score 579.8; DB 10; Length 655;
Best Local Similarity 98.9%; Pred. No. 6.2e-150;
Matches 625; Conservative 0; Mismatches 3; Indels 4; Gaps 4;
QY 2046 CATTGAGGATCACAGGAGATGGCAACTTCATCTATTCCCAAGCTGACGAAATCAGAA 2105

```

[illegible]

```

Db 601 GGTTTTAACCTGGAGTTGGGCTCTCACAGAACATTTCTGCCACTGGAGCATGAC 660
Qy 1999 AATCAGCTGACAGCTCAAGTGGAGTGTGTGACACAGACAGGGACCCATTTCAGGATCAC 2058
Db 661 ACCATGACACAGCTCAAGTGGAGTGTGTGACACAGACAGGGACCCATTTCAGGACCAT 720
Qy 2059 ACAGGAGATGGCA 2071
Db 721 ACAGGAGATGGCA 733

RESULT 14
Bi697844
LOCUS 603343961F1 NCI_GAP_Mam2 Mus musculus cDNA clone IMAGE:5371862 5',
DEFINITION mRNA sequence.
ACCESSION Bi697844
VERSION Bi697844.1 GI:15660473
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 793)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: crapps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M1946 row: 1 column: 15
High quality sequence stop: 748.
FEATURES
Location/Qualifiers
1..793
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5371862"
/clone_lib="NCI_GAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 240 a 176 c 198 g 179 t
ORIGIN

Query Match 20.6%; Score 571; DB 13; Length 793;
Best Local Similarity 84.5%; Pred. No. 1.9e-147;
Matches 665; Conservative 0; Mismatches 120; Indels 2; Gaps 2;

Qy 741 TTACACCGACAGCGGATAGCAAGAAGAGTTTCTCAGCAACTACAGTGTTCGACAG 800
Db 1 TTACACTGACAGCGCAATAGCAAGAAGAGTTTCTCAGCAACTACAGTGTTCACAG 60
Qy 801 CAGTCTCTCAGAGATTTCATGCAAGCTCTGGCATGGATCAGGAGAAATTCAC 860
Db 61 CAGCATCTCTGAAGATTTCATGCAAGCTCTGGCATGGATCAGGAGATTCAC 120
Qy 861 TTCTGACAGATCAGAGCTTCTCCAGTATAGCAACCACTGGTCTGACAGCGCTCCCG 920
Db 121 TTCGATCAGATCAGTGTCTTACAGATGATGGTACCAACTGGTCTGTAGAGCGCTCCG 180
Qy 921 CTTGAACTACCTTGAGATGGTGGACTCCCGGAGAGATTCTCTACCGAGAGTGGATACA 980

```

```

Db 181 CCTGAACCTACCCCTGAAAAATGGGTGGACTCCAGAGAAAGACTCTCTACAAGAGTGGATCCA 240
Qy 981 GGTAGACTTGGGCTTCTGCGCTTTGTCTCAGCGCTGTCTGGGACACAGGGCGCCATTTCACAA 1040
Db 241 GGTGGACTTGGGCTTCTGCGCATTCGTTACTGTGTAGGGACACAGGGTGGCATTTCACA 300
Qy 1041 AGAAACCAAGAAGAAATATTATGTCAAGACTTACAGAGTAGACATCAGCTCCAAGGAGA 1100
Db 301 GAAAACCAAGAAGAAATATTATGTCAAGACTTACAGAGTAGACATCAGCTCCAAGGAGA 360
Qy 1101 AGACTGGATCACCATAAAGAAGGAAACAAACCTGTTCTCTTCAGGGAACACCAACCC 1160
Db 361 GGACTGGATCCTCCCTGAAAGAGGGAATATAAGCCATTATCTTTCAGGGAACACCAACCC 420
Qy 1161 CACAGATGTTGTTGGTTCAGTATTCCCAAAACCACTGATAACTCGATTGTGCCGAATCAA 1220
Db 421 CACAGATGTTGTTGTTAGGAGTTTCTCCAAACCACTGATAACTCGATTGTGCCGAATCAA 480
Qy 1221 GCCTGCAACTTGGGAACTGGCATATCTATGAGATTTCGAAGTATACGGTTGCAAGATAAC 1280
Db 481 ACCTGTATCTCTGGGAACTGGTATATCTATGAGATTTCGAAGTATACGGTTGCAAGATAAC 540
Qy 1281 AGATTATCTCTGCTGCTGGAATGTTGGGTATGGTGTCTGGACTTATTTCTGAC-TCGCCAGA 1339
Db 541 AGATTATCTCTGCTGCTGGAATGTTAGGCACTGTGTCTGGACTTATTTTCAGACATCCCA 600
Qy 1340 TCACATATCAACCAAGGACAGACAGAAACTGGATGCTGAAACAACTCCGCTGGTAAACA 1399
Db 601 TTACAGATCCANTCAAGCCGACAGGAATTTGGATGCCAGAAACATCCGCTGGTGACCA 660
Qy 1400 GTCGCTCTGGTGGGCACTTCCACCCGACCTCACTTACATCAATCAGTGGTTCACAA 1459
Db 661 GTCGTACCGGCTGGGCACTGACACCCCTCAGCCCAACCATACA-CAATGAATGGTCTCAAG 719
Qy 1460 TAGACCTGGGGAGAGAGATCGTGAGGGGCATCATTCACAGGTTGGGAGACACCCAG 1519
Db 720 TGGACCTGGGAGATGAGAAGATAGTAAAGAGTGTATACATGATCAAGGGGGGAAAGCACAG 779
Qy 1520 AGAACAA 1526
Db 780 AGAAAAA 786

RESULT 15
Bi696148
LOCUS MR3-UT0091-150900-005-e03 UT0091 Homo sapiens cDNA, mRNA sequence.
DEFINITION MR3-UT0091-150900-005-e03 UT0091 Homo sapiens cDNA, mRNA sequence.
ACCESSION Bi696148
VERSION Bi696148.1 GI:10901858
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 627)
REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Balag, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202863
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

```

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MR3-UT0091-150
900-005-e03&t3=2000-09-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 26
High quality sequence stop: 564.

FEATURES

Source
1. .627
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0091"
/dev_stage="Adult"
/note="Organ: uterus; tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 187 a 130 c 172 g 138 t
ORIGIN

Query Match	20.6%	Score 570.4	DB 12	Length 627
Best Local Similarity	98.8%	Pred. No. 2.5e-147		
Matches 585	Conservative 0	Mismatches 6	Indels 1	Gaps 1
QY 2181	CATGTCCTGGTCCACGCTCGGCACACTCAGGTCACACTGCCGTACCCAGAGCCAGAGGA	2240		
Db				
QY 2241	GTACGATCAGCTGGTCTGGATGGCCATTGGACACCAAGGTGACCACTGGAAGGAAGGGCG	2300		
Db				
QY 2301	TGTCCTGCTCCACAACTCTCTGAACTTTATCAGGTGATTTTCGAGGGCGAAATCGGAAA	2360		
Db				
QY 2361	AGGAAACCTTGGTGGGATTCGTGGATGACATTTAGTATTATTAACACCATTTTCAACAAG	2420		
Db				
QY 2421	AGATTGTGCAAAACACGACAGCTGGATAAAAGAACCCAGAAATTTAAATTTGATGAAC	2480		
Db				
QY 2481	AGGGACGACCCAGGATACGAGGTGAAGGATGAAGAGTGAACAGACATCTCCAGGAAGCC	2540		
Db				
QY 2541	AGGCAATGTGTTGAAGACCTTAGAACCCATCTCATCACCATCATAGCCATGAGCGCCCT	2600		
Db				
QY 2601	GGGGTCTCTCCCTGGGGCTGTCTGTGGGTCGTGTGTACTGTGCTGTGGCATAATGG	2660		
Db				
QY 2661	GATGTCAGAAAGAACTTCTGCTCCCTGGAGAACTATTAACCTTTGAACCTTTGTGGTGT	2720		
Db				
QY 2721	GAAGTTGAAAAGACAACTGAATACACAGAGTACTTATTCGGAGGCATGA	2772		
Db				
QY 2772	GAAGTTGAAAAGACAACTGAATACACAGAGTACTTATTCGGAGGCATGA	2828		
Db				

Search completed: November 19, 2002, 23:59:40
Job time : 2579 secs